

<110> Folkesson, Anders

<120> The complete sequence of the cs7 insert in Salmonella enteric serovar Typhimurium

<130> Complete sequence of the cs7 insert

<140>

<141>

<160> 5

<170> PatentIn Ver. 2.1

<210> 1

<211> 46870

<212> DNA

<213> Salmonella typhimurium

<220>

<221> CDS

<222> (37368)..(37868)

<223> safA putative fimbrial subunit

<220>

<221> CDS

<222> (37952)..(38689)

<223> safB putative periplasmic chaperone

<220>

<221> CDS

<222> (38713)..(41223)

<223> safC putative outer membrane usher

<220>

<221> CDS

<222> (41245)..(41715)

<223> safD putative fimbrial subunit

<400> 1

gatacaaatc tcaggggtgtt ttatacatc ctgtgaagta aaaaaaacg tatcactgta 60
aaagggatac gggttttttt cgtcttcaag aagttccacc gtctatcgtg gaatctggcg 120
caaatgggcc tacgcctgga tgacgaacag gatattaccg ccacttcttt cactgtcatg 180
gctattttga tccactgac atttaaggcg cggcctcatg gcggtgctta accgggatcg 240
ggacatgttc agcgcagaag cagactgcgt aatgttgata tctactcagat aattacggag 300
aaccgccaga catgcgcac atcactccag ggcacccac ttctccagca actccaccgg 360
gatctcattg atcacctccg agaaccgttt tcccaccagt ctttcagcct ggcgtaacag 420
tgggatggtc gggctactgg gttcactgct ctcaaaccag cgacgaatgc accgcaggcg 480
ttccagcgca tcattacgat cgcgaattgg ccccggttca gcattgtgtg ggagggatat 540
accgggtggc gcgttgacag gcggcacatc ctctgccgt gccggaacgg cattatccat 600

aaccgtgcct gcgggggtctg gcactggcgc tggcgacgga gggatttccg gtgtggttgt 660
ctgtaccacg tcaggcagca gcgctaacaa ctgccgcagg cgtgaaaaat ctggagccag 720
atcgccaat gtctcagcg cccacacttg caggcggtct gcgcttctct gcgcttcgcg 780
aaacgcgcgc agcggtaatg ccccaacgggc ttcgaggtcg gccagttgct ggcggacgga 840
ttcggggggc agcgcacacg ccggggcggg agcggataac gcgcgctcca catcccgtac 900
ctgtaggcgc agggcgccac tgtttgacag cgtaataccg cgaatgtccg ccatcacgcc 960
ttcgtgatcc agcagcgccg ccagggcatt actgdcgcgc agcgacgcac cttcaataga 1020
ttctgcggat gcgccttcgc cggtaagcag ctgtggatgg agtgcacacg accagattac 1080
gctcagttct gccagttgtg tgagcatttc cgccagtcgc tgtgcgccag cctgctggat 1140
gcgactgcgc agcagcagga tcaacacccg gatatactta ctgcgcgtga gcagacggcg 1200
cgcacacagt tcaatttccg gccagttcac ggcttcgggc gtactgacaa aatcaccata 1260
ttgcgcctcc gcctgggggg cggcgcggtt gaacagcagc aggtattccg gatcgtactc 1320
cggatcgggg ccgcagggtt gttccgcgct taccggttta gtcagggaca tgtccatata 1380
attactctca gtgggttaag ccgtgttcag gttcaaaaat catgccggtc accggacggg 1440
catgcatggc ttacactgcc caggccgtcc agccaagccg ctgcgcgcgc ccgatcacccg 1500
caggcggtgc ggcacagggc ttcaactgca gctcaatttc ccagcaatat tcaaaaccga 1560
tgaacgtacg caccagttct gtcagcacgg gcagattggt gccgcggggc agaaaacgca 1620
ggtaatcttc cagcgtgagc gggccgataa ttagccggaa tttgtactgc atatccggtg 1680
cggcctggcc gataagcgcc ccgttgccca gcaccgaaga ctgcgcgggc gtgccagac 1740
gggtgatttc atcgttcgcc accgttatcc agtgcagggc gaattctttc accgcaaaag 1800
gtacgctgaa atagtgcgc agcgtggcgg ccagcccgtc aggattgcgc gattcgcgta 1860
ccagatgggc ggaggctgcc aggcgaacat gatctgacag cgggctttcg gcgctttccc 1920
gtagatcctg ccgctgaga ctggcgatat aaaacgcaaa acggtcgtgt tccggtttgt 1980
ccagcccgcc accagcggac tgggcgctgc gccatgcctg ccagaactgc gtcagccagc 2040
ggtggtgaaa aatattggaa aaatgaacca gcgtcggatc gtgacgactc tctgagcggg 2100
tcagtgccag ctcggtatag tgcagcggca acgggccgtt tggcccccat agtccgaggc 2160
tgtacaggct cagggtcagg cgtccatcct gccagctgac ctgggcgatt tcccgtggcg 2220
caaaggatcat cgtcggcgtc tgtcccagac ggaatttttc catccgtggc ttgccagata 2280
cttgctgcc ggagtatcac agagctgggc atccacgcgc cgcacaggt tcagggaatcc 2340
gtaacgccag ggggttttct gcgcctggtt gagagcgccg gtcatacgtc cccccgttgc 2400

ccggttctga ccggccaggt tggacatgc ccgcgttgca tcgagtgcag tctcatctgc 2460
 gagaaggtat taatggaaac atggcgggca atatagtgtt ccagcaccag accgaacagg 2520
 taaggactga taccggaaaa tccttcttcg tccacggta gttcgcaact gacgccccgg 2580
 ccatagacca acaggccgga gccgggcagg cggcgggtca ccgggggtgt tttgcagcca 2640
 atcaggctgc gcacctggcg cgactgcggg ctgtcgtgag ccgggataaa cagattcagc 2700
 agatcgcgca gcgcctggcc gccgggtgcg tgatccagat cggccagcgg cagataatta 2760
 aacgacaact gccggatcag ccgccaggcc atttcgcgtt cagccagcgg cggctgcggc 2820
 gggcgcgggc gtctgataag acccacgccc gccaccggaa tcgccgcatac tacggtcaga 2880
 tcatcccgcc cattacgtgg aataaggcag ggcagatcgc ggtagtcac cattgccgtg 2940
 acggtgatat ggcgcagatt ttccgggtag ggcgcttcat gctgatcaac cagcgagagg 3000
 aagacttccg agccggtata ggggggttcg gtgccatagc ggcgggcggt ttctgacgag 3060
 cggcgcggtt cagcagcgag tgaaaaataa cgcccggtgt tgccttcgtc attattacgg 3120
 gtgtgataca gcggacgaaa aatcatcttc cgtgtggttt ccgcttccag gccttcgact 3180
 tcctgaacag aaaacacctc gtaatccagc ggacgggtac gatccaccac cagatgctgt 3240
 tccgtcacgc tgtgagtgac ttcaatccgg gtggtggtgc gaggaagcag gttgatcacc 3300
 ggcgtacaga acaggctgaa ctgtgcagcg tccgtctgat gaatcagcca gtccggcggc 3360
 aggcgggttaa gcagtatgac aatttccgcc acattgccct gcaccttttg taaccgggca 3420
 gacaatccgg tcggggtgaa gaagtaaaac cgttccggac aggcgaaaaa ttcatgcagc 3480
 agattatggc cgtgaaacac gttccaggcg agcggtagca gccctgccc tggctccagc 3540
 ccttcgtgcg ccaccgggtg ttgaagattc acattcagtt cgcggtcaaa gtgaccgggt 3600
 tcaccggcca gtgtggcgac ggcgctggta tgtagcagct caaacagggtg tgacgcaatg 3660
 cgttcttcgc cgcagaggta aaagggcagc cgtgccggac cggccagctc gctgaaagtc 3720
 agttccccga aggttcgcag ggtgatgcgc aatgccccgg cgacatgaat attaggcggc 3780
 agatagcggg gcagggcggg catatccggc ggcgcggcgg tcaggcgtac ctctcgtatg 3840
 gacagcggcc acagcgtgac gtccctggctg ctgcgaaact ggcaggcggg attttcgcc 3900
 tccgggatgg gggaaacgaa ccgggtatcg cgcggcacgg tgacccttt cgccaggtcg 3960
 ccttctcgtc tatcgggata cagctttacc actgccatcg atggcgtggg ggtgacgtaa 4020
 ttggggctga cgacttccag taaccgctgt gtgaagcggg gaaactcggc gtcaattttt 4080
 agctgagtgc gggcgctcag aaagctgaac gcctcgatca tgcgttccac atacgggtcg 4140
 gcaatatcgg ttccctgcat cccagtcgg gcggcaattt tgggatggag ggtggcgaac 4200
 tcagcaccgg tctcccgag gtagctcagt tcgcggtgt aatactccag tagccgtgga 4260

tccatgaata atgccctgta ttaaaagaac gtaatgcggc tcagctccat atccagcgcg 4320
ctgcgtacca gaaactccgt gggatacggc tgcgtcagaa tttgtccgcg aatttcaaac 4380
tgtagcgtgt tatagctacc ctgccgggtt ttatcgagca agggcgtagac cctgagtgtg 4440
gcggcggttca gdcgggggttc gaaacggata atggcgcgcc ggatcgccctc gctgatatcg 4500
tcccacttat gctcatacat aaagctgccc gccagcgggc gcaggccata gttgagcact 4560
gacgcgcgcg cctgcgggata gcgcggggcg tcgatgtcac cctcgtggct aatgggtattg 4620
agcaaaaagg agagatcccg gcgaatgac tccttcagtt gtaccggcgt gacgctgata 4680
tcccgggtcaa ttttctgata cggagcattg tcacacagcc gatcaaacag cgtgggcagc 4740
aggtgggttag cgggtgtaaa acgagacgtg ctcatgcgcc atcgttttcc tgagcatgaa 4800
aggtacaatg ggccatgtcc agcaggctga tatcgccgtg gctggtcagc cacactttct 4860
gccccagcgc ccgcacgggtg gtttcgccgg ggccgtcctg ccaggcggtt tccctgcaca 4920
gacgcagggc gtcggatgca ctttccgaac cgctgtaacg ggtaaagagc caggcgccgt 4980
gcgtatcgcc attcaccagg gtgatattaa cgggtttcca cagcagatca gtcaggcgcg 5040
tcggttgccg cgattccagc gagcgtattt gcgaaaacgg cagccagata tacacgcgcg 5100
cggtgactag ctcaagtacc gggccaaggc gggaatcgct gtcgctcgcc cagtcaaattg 5160
cgccgccgtt ccaactgccc cccgtgtctg ttatggcttc cagtgcggta ttacggtgtt 5220
tatcaacctc accggtatcg tcatgacagg cgagtgcgcg cagcagtgac tccacccaaa 5280
cgggctgcgg cagaagaaaa ccgggtcggt gtaccacctg aaaaacgggtg tggcggaaca 5340
tttcgcagcg aaccagctcc cggtagagcc gggcctcctg ggtataattg gcctccatcc 5400
tggcgcatag ctgaagctgg tgtagcgccc gcgaccagtc tccggccaca cacagcaact 5460
gaaacaggct gtggcggcag agcgctttcg ccggattttc ccgaacctgc tgctccgcca 5520
tctgaatccc ctccgcaata gagtattcct gtatcagcgc ggacagggtg gcaggaagcg 5580
tgtagtttt tttcatgggc ggtatttcca tttttctgtg tcggagtgat tcggtagtgg 5640
ctatcgatgc caataatacg atgctcgcg cggttcagat cgggcagtat ctgctatgt 5700
gcggttttcc ccccagttc cggcgagagc agttgcagga tatcgggcat tgattccatc 5760
gccagccagt gcatttcacc ctccaccggtc gtgtccagcg tatccagaat ggcataata 5820
ccgggcgcgc ctgccaccat atcctgtaac gtgtcgggtg cgccttttt atcgtataac 5880
gacgtcaaat cctgggtgac agcgtcttca tttcggggaa acggttttagc acggaggggc 5940
ttttcctgcc cggagggccg caatgcctgc tgatattcct ggtaaagctg gtgtagtgtg 6000
ctatccgctt catggccagt gtaagaggat gcggtttccc ggacgggaat aatatcgaac 6060

gggttctgcg gattgagctg gttgttta aaccagtcca gatccaggta cagcaacg 6120
gattcagatg actgaaccaa ctgaggaaac gaaacatcgg gacgggactc atcattggtg 6180
cgggcaaggc accaggatga cagaccccat tcaatcgtat ctccatcatt cagccgcata 6240
cgggtgatgag gttccatcac ctgctcattg acggcacagc acagttcac agagtgatta 6300
acaatccacc acgcctcttc gtggcgaaca agactgagcg cgatatcgtc tttcccatgt 6360
tgctcaaaaa gcggaatgta gggggcggcg gcagtaaata cgatgtatgt ccctgcggga 6420
tactgtgctg tatccgttat accctgcgac tgtattttcc gtaattccca catgtcttgt 6480
ttcattatgt cccttaaatg tatttatttt tctggaggaa acgtttaggg agttttaatt 6540
cataaataat tttactaaa tttataggga gtcattattg atgacaccct tttattatt 6600
ttcgtcgtga atgcattgtt gtatgcatag atgtcttttt tgaaatatta tttcttttaa 6660
ttctgcaatt gcgattttta tcttgctgtt cattgattat ttaatatata cggcttagtt 6720
gctaattagt ccctgatctg tatcattgtt ttgtttcgat attttttcga ggctatcaat 6780
aaagaattgt ttatatattt atatgcattg atgcattatt tttatgaatt tttatgtcac 6840
aaggcataac acatggaaac tcctgtttca cgcagtgcgt tgtatggaaa actggccggc 6900
ccactattcc ggtcgtgga atcggcaacg gcattttgca aactacgctc taatccctgg 6960
gttgagctga ctactggct gcaccagtta acacagcagc ccgataacga tattctccac 7020
gttcttcggc attaccagat ccctctttct gatgtggaga aagcgttact cgggcaactg 7080
gatatgctgc ccgcccgggc cagcgccatt agtgattttt ctcaaccatat cgatctcagc 7140
gttgaaaagg cctggatgct ggcgagcgtc cgttacggcg ataacaaaat tcgcagcggc 7200
tggttgctgc tggccttggt gaccacgcca gaactgcgtc gggactgag cagtatctgc 7260
gcgcccgtgg ccacgcttcc gggtgatgaa ctgacggaaa tactgccctc gttgatcgaa 7320
acatcgccgg aagcgcagga gcgcccttac gacggctccg gcctggcacc agccattccc 7380
ggtgaaagca gtcaggcgat tcccaacggc gggcaggacg gtaaatacgc gctggcaaaa 7440
tactgtcagg acatgacggc acaggcgcgc gacggcaaaa tcgaccgggt gacggggcgt 7500
gagcatgaaa tccgcaccat gacggatatt ctgctgcgcc gtcgccagaa taatccacta 7560
ctgactggtg aggcgggcgt cgggaaaacg gcggtcgctg aaggttttgc cctcgcgatt 7620
gcgcaggggg aagtgccgcc cgcgctgcgg gaagtacggc tactggcgct ggacgttggc 7680
gctctgttgg ccggagccag catgaaaggc gagtttgaat cgcgtctgaa agggttactg 7740
gaagaggccg ggcgctgcgc gcagccgggtt attctgtttg ttgatgaagt tcacactctg 7800
gtgggcggcg gcggcgcatc cggcacgggc gatgccgcta acctgctgaa accggcgctg 7860
gcgcggcgca ccctgcggac tatcggcgcc accacctgga gcgaatacaa gcgccatatt 7920

gagaaagatc cggcgctgac ccgtcggttt caggtgttgc agattgccga accggaagag 7980
 atccccgcaa tggaaatggt gcgtggtctg gtggatacgc tggaaaaaca ccataacgta 8040
 ctgattctgg atgaggcggg acgtgcggcg gtacagcttt ctcaccgcta cattccccgcc 8100
 cggcagttgc cggataaggc catcagcctg ctggataccg ccgcggcccg cgtggcgctg 8160
 acgctgcaca cgccgcctgc cagcgtacag ttcttgcgcc agcagctaaa agcggcggaa 8220
 atggaacggt cgctgttgca gaagcaggaa aaaatgggga ttcagtcaga tgagcggcgc 8280
 gatgcgctga tggcgcgaaat tttctcgctc aacaatgaac tgactgcac cgaatccccgc 8340
 tggcagcggg agctggaact ggtacatacg ttgcaggaaac tgcgtctcgc agagtctgat 8400
 gctgatgaca aaaccacgct gcaacaggcc gaaacggcgc taaggagtg gcaggcgac 8460
 gcgccggtgg tgttccccga agtcagcgcg gcggttgctg cggcgattgt cgcgcactgg 8520
 accggtattc ctgctgggcg catggtgaaa gatgaggcca gccaggtgct ggaactgcct 8580
 gcccgactgg cgcaacgctg taccgggcaa gacggcgcgc tggcgagat tggatgaacgt 8640
 attcagaccg ccaggcgggg actgggcgat ccacgaaac cggtgggcgt gtttatgctg 8700
 gccgggcccgt ccggtgtcgg taaaaccgaa accgcgctgg cgctggcgga ggctatctac 8760
 ggcggcgagc agaacctggt aaccatcaat atgagcgagt tccaggaggc tcacaccgtt 8820
 tccacgctga aaggcgcgcc gcccggtat gtgggctatg gcgagggtgg tgtgctgacg 8880
 gaagcgggtg gtcgccaccc ctggagcgta gtgctgctcg acgagatcga aaaagcgac 8940
 catgacgtcc acgaactctt ctatcaggtg tttgacaagg gtgggatgga ggacggcgag 9000
 ggtacacatg tcgatttcaa aaacaccacg ctattactca ccaccaacgt gggttccgac 9060
 ctcacagcc agatgtgtga agatccggcc ttaatgccc atgctacggg gcttaaagag 9120
 gcgctaattg cggaattgcg caagcatttc ccggcgcat ttctgggccg cgtgacggtg 9180
 atcccttacc tgccgctgga cgaaacgtcg cgtggcgtga ttgccgtct gcacctgac 9240
 cggctggtgg cgcggatgag tgaacagcac ggcgtgacgc tgacgtatag cgaggaactg 9300
 gtcgcacata ttgtggcggt ctgtccaatg catgaaacgg gcgcgcggtt gctgattggc 9360
 tacatcgaa agcacattct gccacgactg tcgcgtact ggttgcaggc catgacggaa 9420
 aaagccgcga tcaggcagat tgatatcggc gttaatggtg atgagcagat tgtttttgag 9480
 atcgtttctt gaaaccggcc gttcgaagtg tccgtagtgc gattttaaaa actgtaccgg 9540
 tataccgctc cccttgcggc aaccagttga ctaaaaagga aatgaaggat tatggctatc 9600
 aacaatagcg cgcagaaatt catcgcgcg caccgcgcgc cgcgcgtgca gattgaatat 9660
 gacgtagaga tttacgggtc cgagaaaaaa atcgagctgc cgttcgtgat ggcgggtgctg 9720

gccgatctgg ccgggaaacc gtgaagaa ctgccgccgg tgacggatcg aaattcttc 9780
gatattgata ttgataactt caatgagcgc atgaaagcca ttgccgccgg cgtggcggttc 9840
gctgtgccga atacgctgac ggggtgaaggt cagttgatgg tcgatatac gctggaaaat 9900
atggacgact ttctgccggc acagattgcc cgcaagggtg acgccctgaa ccagttactg 9960
gaagcccgcga ctcaactggc gaacctccag acctacatgg atggcaaggc gggggcgga 10020
aatctgggtca ataaactgtt gcaggacccg actctgctga aaacgctggc gaatgcgccg 10080
aaatcggccg ctaccagca agatgtgtca gcggataatg aatcagcga ataacgtcga 10140
atttttaagg aattttcatg gcaaacagta atatgcaggc aaccgacgcg gttgctcagg 10200
ataccgcctc cgcattccgt gaatttgatg cgttgctgaa tcaggccttc cgaccaaga 10260
ctaccagggc ggcaaaagcc gtggaagccg cgggtgcagac gctggcgaac acgatcaccg 10320
tcagcgatga cgcctataaa agcatcagcg ctattattgc gcagatcgac tttaaactga 10380
ccgaacagat caaactgatc ctgcaacatc ccgactggca gaagctggaa tcctcgtggc 10440
gcggtatgga gcatctggtt tacaacaccg agaccgacga aaagctgaaa attcgttca 10500
tgaatctgtc aaaagatgaa ttgcggcgca acatgaagcg ttacaaggcg atcgccctggg 10560
atcaaagccc gatgttcaag aaactgtatg aagccgaata cggccagtta ggtggcgaac 10620
cttatggctg tatcattgcg gattactact tcgaccatac accgcccgat gtggatctgc 10680
ttggctctat cgccaaagtc gccgcgtcgg cccatgcgcc gtttattgcc ggggcttccc 10740
cctcggtact gcaaatggac tcctggcagg aactggcgaa tccccgcgac ctgacaaaa 10800
tcgtcaccca gaacctggaa tatgcgccgt ggaactcgct gcgggctagc gaagactccc 10860
gttatattgg cctgacgatg ccgcgttttc ttgccgcct gccgtatggc gcaaaaacca 10920
acccggtgga cgagttgat tttgaagaag atgcggatgg ttctgaccat accaaatagc 10980
tctggagcaa cgcggcctac gcgatggcg taaacatcaa ccgttccttc aaacactacg 11040
gctggtgtac gttgattcgc ggtgtggaat caggcgggtgc ggtggaaaat cttccctgcc 11100
ataccttccc gactgacgat ggcgccgtgg acatgaaatg cccgaccgaa atcgccatct 11160
ctgaccgccg cgaggctgaa ctggcgaaaa acggttttat cccgttgatc caccgtaaaa 11220
actcagacta tgccgccttt atcggcgcac agtcgctgca aaaaccacag gaatactacg 11280
atccggacgc gacggccaac gctaacctgt ctgccgtct accgtacctg ttgcctgct 11340
cgcgcttcgc tcaactcttc aaatgtatcg tccgcgacaa aatcggttcc tttaaagagc 11400
gtgaggatat gcagcgctgg ctaaatgaat ggattatgaa ttatgtcgac gccgatccgg 11460
tgaactctc gcaagaaact aaagcccgtc gtccgctggc tgccgctgaa gtagtggtgg 11520
aagaggtcga aggcaatcca ggttattacg acgcgaaatt cttcctgcgt ccgcatttcc 11580

agcttgaagg gctgacggga tcgctgcgcc tggtgacaaa actgccgtca gtgaagcagg 11640
gcaatgcctg atatatat tgtgaatggt taagcgagtg aagtcagaga agatagagaa 11700
tataaagagg gatatgaaga aaagaatttc gtctcgccca cggctctgta aagggtgggt 11760
acgtaatgat gacacatatt cgaatgccag taacaatgcc gaagcttttt atattcattga 11820
gtaggaaata catattatga ccataagccc aacttttcat ctgttacctg gtattgttct 11880
gtctttttca caatatgctg tagcctggga agtcagttgc ccggctgta ttgatactca 11940
gtcttctgct gtgagcctga agtctgatgt ccagcggcg tggcagcttt cccccgata 12000
tatgtcgcgt ttatggttaa gtagtattgg ggtaacgcag ggtaaacctg aaaacctgat 12060
ggatctcaaa ccagagacta aaaaagtaaa cggtgaaaat tggctctgtat gggaaacaga 12120
acgtggtagc gataaagaaa ccgatcgcta ttgggtttcg tgtatttatg gtcattgaaca 12180
gatattggtg acgcaaccaa tacctgcttc ttctactcgc tgtaagactc gtaattttga 12240
gggatcgcca gaagaccagt ctgtatcttt tatctgtaat tagcgatttg agacgtgaaa 12300
atctcagtag aggttatggt ttttattatc ggaagttatg aagcattatt tatatgcatt 12360
aaataatgca aattcataaa ataactaaat acattatcgg taccggaaaa atatacagtc 12420
ctctgttctc ctgaagttat tggagaagga ttctgtacgg caatgattta tctataaaca 12480
aaaagatata gataaaatca gggtttat tttt aagtaaaact taataaggat ataaaaatgg 12540
cttatgacat ttttttgaat attgacggca ttgatggcga gtcaatggat gacaaacaca 12600
aaaatgaaat tgaagtactg agctggcgct ggaatattca tcaggaatcc accatgcacg 12660
ccggtagcgg cctcggtccc ggtaaggctc ccgtcaccaa cctggatttt gatcactata 12720
tcgaccgcgc cagcccgaaac ctgttcaaat actgcgcctc cggcaagcac attccgcagg 12780
ccattctggt tatgcgtaag gctggcgcca atccgctgga gtacctcaag tataccttca 12840
ccgacctgat tgtcgccgtg gtttccccga gcggcagcca cgatggtgaa atcgccctcc 12900
gtgaaacggg ggagctctcc ttcagcaccg tgaagcagga atacgtggtg cagaaccagc 12960
agggcggcag cggcggcacc atcaccgcag gctacgactt caaggccaac aaagaaattt 13020
aacggctggt tttccggcca gatgttatgt ctggctgggt ttattgtttt gattttaaag 13080
gaatttacag tgaataaatg gcgtaacccc actgggtggt tatgtgcggt agctatgcct 13140
tttgactgc tctgcttcc cggatgcggc agtagcgatt cgctacttga cccctaactg 13200
cagcggcctg gcctgagcgt gaaagcgttt tacaagggtga attctgacaa tcagaagaaa 13260
gcggcgcca tgaagatacg tgttgagaat taatgacctc cacagaattt ttagagggtta 13320
agcaaaatga acagaccttc attcaatgaa gcgtgggttag cttttaggaa ggtgaatcat 13380

tccgtcgctg atgtgggtag ttatttgggt ggaaacgttg ggaaaaatat aactgggtggt 13440
tattttcaaa atgcctgccg tattcgaatg agctatgttt tgaatgcgac agggttccca 13500
atagcccgtg actctccgta tgcaaagggt agtggtgccg ataataaatt ctatatttat 13560
cgcgatgaatg atatgattga ttatcttact catactatgg gcaagcctga tcttattggt 13620
aataatccga aacagagtga ctttatcggg aagaaaggaa ttatcgtagt aaaagggcat 13680
ggctggagca atgccagagg acacgttaca ttatggaatg gcagtatctg ttcagatcag 13740
tgccacttat taaatgaccc agataatgga ccatttggtc ctgaagtggg gacactgtgg 13800
atactgccgt gaaatgggta atattagtta ctttcagcat aagtggaaatg ctggatatggc 13860
agccatcctt tgcacaagaa gcattgacca cacaatattc acagtcggaa ctcttaaaaa 13920
attgggcgct gagtcattgt ctggcattag tatacaaaga tgatgtcggt aaaaacgatg 13980
ccagagctac ggccagtgtc taccttgaat atggtaaaca atctgtggag atttaccatg 14040
aaattgatga gattgcgaaa aaatattcag ggttgaaata taacgggttcg atatcatcag 14100
attttaatac catgaagtgc atagatttta tccatgacag ggaattaaat gaattaatta 14160
aaaggcggtg cgagaagtaa aattcaaaga tattaaaagc atacgttctc ttgctctgat 14220
gaatttatgg gtaagaaaga gctgtacagg aatagttaat ctgttcacct aataaagcag 14280
ataaatcagg gcttaattta ggtagttaa aggatagtag atatgtctta tgacattttt 14340
ctgaaaattg acggcattga cggcgagtca atggatgaca aacacaaaaa tgaaattgaa 14400
gtactgagct ggcgctggaa tattcatcag gaatccacca tgcacgccgg tagcggctctc 14460
ggttccggta aggtctccgt cactaatctt tcatattgaa attacatcga tcgcgccagc 14520
ccgaacctgt tcaaatactg ctcttccggt aagcacattc cgcaggccat tctggttatg 14580
cgtaaggctg gcggcaatcc gctggagtac ctcaagtaca ctttcacaga tctgattatt 14640
gcaatgggat cgcccagcgg aagccaggga ggggaaattg cgtctcgca atcaattgaa 14700
ctctccttca gcaccgtgaa gcaggaatac gtgggtgcaga accagcaggg tggcagcggc 14760
ggcaccatca cgcaggtta cgacttcaag gccacaaaag aaatttaacg gctgtttttc 14820
cggccagatt tatatctggc cggatttatt attttgattt taaaggaatt tacagtgaat 14880
gaatggcgta accccactcg gtggttatgt gcggtagcta tgccctttgc actgctcctg 14940
ctttccggat gcggcagtag cgatgcgtta cctgacctcg aatcacagcg actcgacctg 15000
agcgtgaaag cctccgataa ggtgaatcct gacaatcaga agaaggccgc gccattgag 15060
atacgtgttt atgaactgaa aaatgacgcc gctttcacga cagctgatta ctggctcgctc 15120
catgacaacg acaaatccgt ccttaccgac gatttagtgc gtcgcgacag ctttattttg 15180
cgtcccggcg aagagaaaaa actgcgtcgc ccgctgaatg cgcagaccac ggcaatcggc 15240

gtactggccg gataccgtaa cctggccaaa tcggtctggc gggtaacctt caaaatcccc 15300
 gaagccccgg aaaaagcctg gtacagcagc ttcatactcg ggaaggaaa agtgcagttg 15360
 gaggcggaac tggaacaaag cgccattgta attacggaac gggataaatg aattatgagc 15420
 tggaatgacc gcgtagtctg gagtgaagga caatctttac tgccgcagat gtttcagcag 15480
 caagagcggt atctggaaca cgtcatgcat taccgcagcc tgccgctgac cccctttttc 15540
 tggggattca gccactacaa tattgatggc gaagcgctga acatcggtta actgatactg 15600
 aaagaggcat cagggatttt tcctgacggc acgcccgtta acgcaccgga ccacaccccc 15660
 ctgcccgcgc cactgaccat tctgcccggg cacctgaacc agcagatttg tctggcggta 15720
 ccggtacgcg cgccgaacag cgaagaaacc acgtttgaca ataaccgga atcattggcg 15780
 cgtttctcgg tacatgaaca cgacatccgc gacgccaact cgctgggacg tggcgcgagc 15840
 ttattacagc tcagtcactc gcgcctgcgg ctgctgcccg aaaaggcggg gacgggcgcc 15900
 tggattggcc tgccgttgac ccgcataccc gggttgaacc ctgacgggcg gatagatctc 15960
 gaccacgacc tgatcccgc catcataat tatcaggcca gttaactgat gtgtacctgg 16020
 ctgtcgtgga tcaacgatct catccggatg cgggcccatt cgctggcgga acggctgacc 16080
 ggcagcgaca accacggcca tgaagcagcg gaggtctccg attacctgct gctgcaaatt 16140
 ctcaatcgct ttgagccgct gctgactcac ctggcgaaaa ccccgctggc cccggagggt 16200
 ctgtaccgct acctgtccga actggccggg gaactctcca cctatgtgcg tccacaaacg 16260
 cgacggcccc ctgaatacaa agagtacaaa cacctgacgc cctatgccg gttgaaatcg 16320
 ctgggtgatg aggtgcagtt cctgctgaac gcgggtactg tccggggcgc gcagcgcctc 16380
 gagctgaaag aggggactta cggcatacct aatgcgggtg tggccccctc cgatcttgcc 16440
 gatttcagca cgctgggtact ggcgataaag gcttcaatgc cgaccgatgt gctactgcaa 16500
 cattttgccg ccagaccaa aatcgggcca tccgatcgcc tgccggaact gatccgctcg 16560
 catctgccgg ggctggcttt gcaggttctg cctgtaccac cgcgcaaat cccgtttcag 16620
 gccggataca tctattacga catccgccgc gagggagcat tgtgggaaca cattgcccgt 16680
 tacggcggga tggccatgca taccgccggg gaatttcggg ggctggagac agaactgtgg 16740
 ggagtgcgag ataaatgaca gacagtaccc tgacgccgcc agcggcggtat atgatgtcct 16800
 ttttgtccac cagccgggaa cataaggaca gtgaatatga aacccgggta cacaccagcc 16860
 agcgcacgga actcaatgtc atcgttgaag acggtccgga cagcaaactc cggctggctg 16920
 aaatcagcgc ggcggctaac ccgttgctcg ccgctgcccg gcctttattg tgcgtctcg 16980
 cagccatgcc cgctaaactg gatgcggccc tggtagagcc ttaccgtaac ctgctggtac 17040

gcgagatgca tctgtaccag attatgcg atcagggcgaa cctgcggcgc ggcacgtac 17100
tggcggtacg ttactgcctg tgtacggcgc ttgatgaagc cgccaataac acaacctggg 17160
gacggcgcg cgctctgggcc ggaaaaagcc tgctggtaac atttcatggt gaaagcgaag 17220
gcgggataaa acttttccag atcatcgggc gtctggcggc cagcttccag gagcatggca 17280
acgtactgga gggtatctac cacctgctgg gggtgggatt tgaaggccgc tacagcgtgc 17340
agccagacgg gcgtaagcaa ctggacaata ttgccagca actgctgaca cagctttcac 17400
agcgtcgca tccggttatg cccgcgctct cgctgactt tcagggggcg ataagcggac 17460
gactgcggcg gatgcgccg gtgccggctt ggctgagcgc cgggatagcc ctgttggcga 17520
tgctgacgct gtttggcctt tacagccacc ggatggatgt gcagaccgtc accgtacaac 17580
agcatattga tgcgattggt ataaaactgc cgccgccgc tgtgccggtt cataagctgc 17640
ggctgaaaat cctgctggca aacgaaatcg cccgtggcct gctgaccgtg gacgaagatg 17700
accagcacag taggggtggtc ttccgtggcg acgcatggt tgtgccggga cagaaaacgg 17760
tgagtacgc aatccggcca gtgattaaca aagcggcgcg ggaaatcgcc cgcgtggcg 17820
gcgcagtcac tgtaacgggt cactgaca gccagccat tcattcggct gaattcccat 17880
ccaacctggt actgtcgga aaacggcgcg cggaagtgc ggcgttgctg acctccggcg 17940
gcgtacctgc cggacgggta catatcgctc gcaagggcga tacggtgccg gtggcggata 18000
acggcagtaa agccggcgcg gcgaaaaacc gtcgggtgga aattctggta gtggagtga 18060
tgaatgatga aaaaatcaac ctatgatgt tctcatcatt cggcagtatg tggcgtgacg 18120
ggggattatt atcggatctc agcgacatat cacataacac gatctgttcg tgttttttg 18180
atcatcttat gttgcctgtt atccgggtggc gcttttgccg gatccccgat taacgcagga 18240
ttcatttccc ccgataatgt caacctcagt actcaggatt tccatgaaatt ttatgccact 18300
gacaacgtac agaaaaaaga caatgcactg atgtatatgc tgggggtgga ggatgcgaca 18360
gaaggtaaag cctggtgtgg atatggtcag gttgacagta taacaataaa ccatactgtg 18420
ctgacctggt ttgaacagca cgcagtga aaagcctgat taagggcttc aatactaata 18480
gaggaagcat tagttaaaaa ttttccctgt cagaggacag actcctccat aaaaattgct 18540
tcccggctcat ctcccatttt atccctgacg ccggatgcgc ttaatctttc aggtaatgac 18600
ttttttaaat tttgggtgtc tggtaatcaa cgggataaac tcagggcggg tgtctatctg 18660
ctcggcgtgg aggatgcgac agagaacaaa ctgtggtgtg gatacgcttt atttaagacg 18720
ctaacattaa atgaattagt ctatgtttct cttaaaaata aaaccaatga ggaactgaat 18780
tctcgcgcg ctgaacttat cataaataaa ttaatagagt atccctgtaa tatataaaat 18840
cattcaagtt gcatcaaggc ggcaaggag tgaatccccg ggagcgtaca ttagttcgtg 18900

actggggtga gcgaggaaag ccaacgcaca tgcagcctga agtatgacag gtataccctg 18960
tgataatggt atcgctgcda gttcagataa agtcttgggt gtaaagtctt ggtggtaaga 19020
tattgaatag ttaaaggtaa tggctgtctt gttgggtgggt gttcctgccca tttcttatgc 19080
gcacgattat ggggtgtgcta cgggttgagc atcaatggag tcatcattat ttgatgccat 19140
aaaaaatgac ctgaatatag atgtcgctac tattataaag gataaaacaa aggtagagat 19200
tcttgatata tcacctgtat ctaaggctta tgcagaatct ctggccagga tggattatga 19260
aaaagataag gccaaaaata aagtggcaat attagataaa aaatcctatt tcgatagtta 19320
ctatgaaaat cagggttaaat ctatcgtggc aaaatatacc tatattaaca aagataaaga 19380
aaaagatatt ttcattgcat ccagcttcat gaatgctgat gagtgttctg taagatttaa 19440
tggttatatt actttatcca gggaatttta aaatattgat tattgggtgg cgagtcgaaa 19500
tataatgcaa aaatttttta gtctgctttt ttcccgccgc gcgctggcag ttgtgggcgt 19560
tctggttctg gcgctgctgg tctggtttgt cgggcccgtg gtgtcatttg ataccctgcg 19620
cccgtggcc tccgtgggta gccgggtagt gaccattgcc ctgttgctga tgcctgctgg 19680
actgtggctg gtcaactggc cgatgagtat catcggcctc agtgtcctgt gcctggcgat 19740
tggcttcgtc acaccgctgc tggccctggg cgatgtccat ccgtttgcgc cgctgtgggt 19800
ccgcctgacc ctgattgggt tcatcctgct gatgtacgcg ctgtacggcc tgtaccggct 19860
gtggcgctgc ctgcgtatgg atgaacaact gctgcgtcgc ttctgcctc gcgcggggga 19920
agaggtaccg gtggcaggcg agatcaaagc cgacctgcgc accgtcaacc atattgtcac 19980
gcaggccatc cggcagctgc ggcagttgcg ggtggatatg cctggctggc gtaaaatctt 20040
cgagggaaaa cgctttctgt atgagctgcc gtggttcatg gtggtcggca gtcccggcga 20100
cggcaaaacc acggccctgc tgaacaccgg attgcagttc ccgctggcgg agcaaatgga 20160
gcagacttcg cgcactctga cagtaccggg tggcggcacg ctacactgcg actggtgggt 20220
taccaacgaa gcggtgttga ttgataccgc cggacgtac gcccgccacg atgacggtgg 20280
tgaagcgagc gccgcgcagc gtaacgcgg agagtggcag ggctttctcg gtctgctgcg 20340
taaacatgc cccggcgcgc cgcttaacgg cgtgatcctg acgctaaacg tggcggattt 20400
aaccgcacag tcaccggcgg aacgcctggc ggctgcgc gctctgcggg cgcgactggc 20460
agaactgcgc gagaccctgg ggattcgtt tccggtctat ctggtggtca ccaaaatgga 20520
tttgttgccg ggggtcagcg aatattttcg cacgctgacc agccatcttc gtgcacaaat 20580
ctggggcttc acgttgccgt acagccgcag gcgaaaagcg ggcgaccgc aggcgctgca 20640
cgccgcctgc gcgcaggagc tggcgcgcct gacgctgcgg ttggatcagg gactggatac 20700

ccggttacag gaagagtagc ctttaaaag ccgccagcgg ctgtatacc cccgcgtga 20760
gttcgccgcc ctccggcagc cgttgctgga ggctattgaa cagatcttcc tcgattcaaa 20820
attcgatgcc acgcaactga ataacacgct gcgcgggggtg tttttcacca gcgccgcgca 20880
ggcgcaggcc gacgccgtgg ccgaccagtt gagtatctgg cagcgctttg tccgggcgat 20940
aaaaaccgcc cgtggcgaat cctccgcctc tctccacac gctctgccgg acggcaaccg 21000
cagctacttc ctgcatgacc tgctgacaca gtttattttt cgtgaagcgc acctgggtgga 21060
gccaaacctc cagtgggcct ggcgttaccg cctgctgcgc ctccggcggc acctgctggt 21120
actggtgctg gcattcctgt tgtggcaggg gatgcagacc agccagcaga ccaacggcga 21180
ctatctgaat gaaatcagcg cccgcgcgac ccggctggac ggtgatgtga aagcctacac 21240
cggtaaaccg gcgatggctc ccgtcccggc actgctggac agcgcaaggg aactgtccgc 21300
ctggccggaa ctggaccggc acgcgccgcc gctggcctgg cgctacggtc tgtacagcgt 21360
accgccggta accgacagcg tggcgctcgt gtacaaccgt ctgctggatc aactgctgct 21420
gcgcgcgctg gtgaaacgga tggagtatgt gctggcggac gccattgccc gtcaggatag 21480
taaagcggcc tacgatgcc tgccatctc tctgctactg aatctggata aagatcacga 21540
agataaatac aacgcggcgg agatccagtc gtgggtgatt aacgatctgg ggaacagcga 21600
cagcgtggcc ggggtccggc ggcgcgcgc cgtgctgacg catatcgaag cgctgtttga 21660
cggcagccgg gtggtgcatt caccgtatga gaaagatgag gcgctgatcc gccaggcgcg 21720
ggcattcctc gacggtcaca ccagtaccga gcgtatctac gcgcgggcgc tggcggcaat 21780
ggagagcgaa gcgcgcagg agttcacgct ggtacgcgcc gtcggcgcgg atgcgggaac 21840
ggtctttgtg cgtagcaacg gcgcgcgct ggatcggggc gtgccgggta tttttaccgc 21900
tgaaggatac cgggagctgt tcgacaaacg attaccggaa tttgtggcgg cggcgacggc 21960
gaacgatggc tgggtgatgg gccgggagag tacgccaaaa aagctgactg acagcctgcg 22020
cagccagata ccggggcagg agcagtctgt cgcgcgcgaa gtccgccgtt tgtacctgac 22080
ggaatatgcc cgcgcgtggc aggattttct ggacagtatc catagtatca acagtgccgg 22140
ggaagagggc agttccggcc tggcctatga tttacagggtg ctgcgcaccc tggcgtcgcc 22200
ggactcaccg ctgatgcggc tgggaaaagc ggtggtggag cagaccacgc tggtgccgcc 22260
gccggaccgc caggccagac agaaacaact ggcgcagcgc gcatccggta acgcggggaa 22320
agtgttacag acggcaaaac tgttcagga tattcaccgc gaagaacggc tggaaaaaac 22380
gctggtggat gatcgtttcg ccgcgctgcg cgaggtcatt gccgggcgga cggacggcgg 22440
acagagcggc ggtgggacga tgcagatgc ttcgctgctg accatgctca acgagtatta 22500
caccctaactg accattgctg atagcgcgct ggcggcgggg acgttgccag cgcgcattac 22560

tgccgcagac aagttgcaac tggaggcggc gaaactgccc gcgccgctga aaaacatcct 22620
gctggatctg acgaagcagg gaacgcgcaa aatcaacgcc gggaccggcg acgtgctgaa 22680
caccagatg gaggcgatga tgggcgacga ctgccgtgac gccatcgacg ggcgctatcc 22740
gttcgccgac agtccgcagg aggtcagcgc cgaggacttt aaccgcatct tcgccagcgg 22800
cggcgctactg gatgctttct ggagcaaaca actggctccg ctggcggaca ccgccagcga 22860
cccgtggcgc tacaaaccga ccgaaggcaa catgacattg cagggggccgg atctgacgcc 22920
gtttcagcag gcgaagcaga tccgcagcgt attttttaac agcgagggcg ggaaaaaatt 22980
ttcctggctg atgcagatta gcgtgggtga tatggacccg gccatcacgg aactggtgat 23040
tgatatcgac ggccagggtg tgcgctacgc ccatgggtccg gaccgtcccc tgaaagtaac 23100
gtggccggga ccgcgcaacg gctcgatggc ggaaatcacc gccagcccgc gtattcgcca 23160
ggacacctca acgctactga ccggcggggc gtgggcgctg tttcatctgc tggacgccgg 23220
aatggtacag gaaacggcgg tgcgggggcg tcaactggtg gaatatgatt ttgacggctg 23280
ccgggtggtg ctggaaatca ccgccgggcg ggattttaac ccggtcagcc gggagctgtt 23340
gcagaacttc agttgcccgg cgagggcgct gtaatgcgc gccagcagc ggtatcagcc 23400
ccggcgatgt tcggcaggct gccggaccag cgcgactatg tgcgctggcg agtgggcgct 23460
gaagaggggac gaatatggca ggactggctg aaccgccaga cgtgggtggg cagcgggcgg 23520
catatcgtgt tgccgggagg ccagacgacc gatgaagagc gtgacggctg gatgcacctg 23580
tcgccgcgtg gcgaacatga tatgcccgcg ccggaacacc atccgctgcc gtggagcttt 23640
gtgatgtcgc cgggattttt gccgatgggt gggacgactg actggctgac cggagtgtg 23700
atggcgtcac gggatagcgt gggacgcccg tggccgctgg tgatttacca gcgctgcgg 23760
cgggagtggc tggacgagag tttgcaggag acgcagggct ggctgtactg gctggcgagg 23820
ctggcggcgc agcatatcac gccggacacg atgcggcgcg ggcggctgac cgagcaggtg 23880
gaccagctat gggcgatgtg gcagccgggg ccgtgggtggg cgcagtggct gcgtgggtta 23940
cgacgtacat cgcaacgcag ccgggaactg accgggctgc cggatgaagc accggtagta 24000
gaactgccgg gcgtgcgtta tctgccgtgg ccgggctggc cgggtaagac gctggggcag 24060
gcaacaccgg ggcaaggctg gttctggcag cagaacagtg aagggcggtg tgtggatgcg 24120
ctgcggttg tggatgaaaa aatagaggt cagatgctgt gaaaaagtac gctgtagagg 24180
ttctgtttat gtctgcctgt gcaggaatgt tcctgcccgt atttgcttgg ggcggaacag 24240
atgtaaacad tgataaccg ctggcggagt gtgtggatat ccatcccgtt catcgtcagg 24300
aatggataa ccttaccatt ctgaaaacaa cggtcacact gaaaaaagc accggtgaat 24360

gtggctgttt ctcaacatta ccaactata ccagcttact ggcgaggat gttgagggtt 24420
 acggacgcgg aagtgcata tcccttcagg aggggaatat ctactggcg aagatgcagg 24480
 gacgttatcc tttagcttt gtattatcgg tggataacca gtctgtacgg gatcagaagc 24540
 tggcgctaata gatacgttgt acaccaccgc tgaatacac agaatagtca gggagaagat 24600
 gatggcagta agactgactt ttgacgggca aaagctgaca tggcctggta tgggatatt 24660
 taaggcgacc acgggggttac cggatttaca gtggccagat aaacagtgtg tgccggatgc 24720
 ggcgataccg gaagggaatt ataaattgtt tattcagttt cagggggagg caccgataag 24780
 aaatgctgcg gattgtgatc tgggaccatc atggggctgg agtaccattc cgcgaggcca 24840
 ggctgccgga acatgtgaga tatactgggc gaactgggga tataatcgta tccggctgga 24900
 atcagcggat gagaagacc gaaaagcctg tgggggcaag cggggtgggt tttatatcca 24960
 tgattccacc aaagggttaca gtcattggtg tattgaagtg gaacgggtgt tttccgtat 25020
 tctgaaacag gagacggaaa aagaaaatgg tgaagagaca ttacgggtta atgttaagta 25080
 tgtttctggt cagcaaacga atggtggaac aaaacaataa taccgttaa accggatgaa 25140
 attatcgtgt ctggacgtgc tgtggtgac aatgattaca atggtggcca gataagtaat 25200
 gagaccaat gataaagtat atgactggcg gtgtgctggt tattaatcaa atttctataa 25260
 aatgcaatgc gaatctgga agcgtataa aaataaaatc ataattattgt tatgttcatt 25320
 tccttattta tgtaattcag ttttatgtt atgtgcta ttttgtgtt ttattttcat 25380
 ggctctgtc agcaatatac cctgttcttc tggtaaataa ttaaataa caggctgggt 25440
 gcattataaa gtgcggggca ctgttctcct acggtgagtc ttttatttt aatccggtat 25500
 taaaggagtc actaccatga gttttgtatc cacaataat aaatccggtg tgggagggt 25560
 gacgacaacc acgccgcga taaccggaga aagtggcgt gtcaccgcag attcagtcgc 25620
 cggagcgtg gcagatgcg cggaatccgc cgtggaacag gctgcgggat cgctatttgg 25680
 cgcattgccg gagccatcag gactggtgaa agccgcggtg gcagcggcg aggctgccgc 25740
 cgccgcaggt atggcgagg atgcggtatc ggccatcgtc tctgctgtt caggcgggccc 25800
 gggggcgcat aatgtgacgg tcagcggcag cgccgtaccg ccgggcgcat tactgttcgc 25860
 cagcctggac ggcggcgaaa cattaagtga actgttcagc tatgtggtac agctaaaaac 25920
 gcccgacacc ctgaatctgg gctatgtctc cccggcggcc aacctgccgc tcaaaccgat 25980
 ggtgggcaaa gatctgtgcg tcaacatcga actggatggt ggcggtaaac gacatatcag 26040
 cgggctggtc acggcggcg ggggtggtgg ccatgaaggg cgttcgggtta cctatgagct 26100
 gcgtatggag ccgtgggtaa aactgctgac ccataccagc gactacaaag cattccagaa 26160
 taaaaccgtg gtggatattc tggatgaggt tctggcgga tatccctacc cgggtggaaa 26220

gcggctggtg gaaagctacc cggtagcac ctggcaggtg cagtagcgtg aaactgattt 26280
 tgattttctt cagcgactga tgcaggagtg gggcatctac tggtaggttg agcacagcga 26340
 ggacagccac acgctggtgc tggcggatgc catcagcgcc cacaaagcat gtccggactc 26400
 gccgctggtc gagtggcacc aggaagggct gaagctggac aaggagtta tccacactat 26460
 cacggcaaac gagagcctgc ggactggaca gtgggtgctg gatgatttcg attttacgaa 26520
 gccacgttca ttgctggcaa acaccgtggc aaaccgcgt gaaaccggtc atgccaccta 26580
 cgagcattat gagtggccgg gagactactt cgacaagagt gaaggcgaga tgctgacgcg 26640
 cattcgatg gaagcgcagc gcagccccgg cagtcgggtg ctggggggag ggaatatccg 26700
 cacactcatg accggttata cttcacgct ggaaaactat cccaccgccg aagtcaatca 26760
 ggaatatctg ctgatgcaga cttgctggt tgtgcaggac aacgcgcagc acagcgggca 26820
 ggaccagcac ttacctttt ccaccggtt tgaactgcac cccaccgcg aggtgttccg 26880
 cccgcagcgg acggtgagca aacccacac caaagggccg cagagcgcca tcgtcaccgg 26940
 cccggcgggc caggaaatct ggacggatca gtacgggcgg gtaaaggtag agtttggttg 27000
 ggatcgctac ggcaaatg atgaaaacag cacctgctg atacgcgtca gctaccgctg 27060
 ggcgggcaaa ggcttcggga tgatccagat cccgcgtatc ggccaggaag tgctggtgga 27120
 tttcaaaaac ggcgatccgg atctgccgat catcgtgggg cgtacctaca accaggacac 27180
 catgcgcgcg tggggactgc cgggaatggc gtcgcagagc gggatcttca gccactcgt 27240
 gtatggcggg ccaacgaacg gcaacatgct gcgttttgac gacaaaacgg gcgcggagga 27300
 agtgaagttc cacgcggaaa aagatctcaa caccacggtg aagaataatg aaacgcatac 27360
 ggttatggtg gatcgacta aaaccattat taaaaatgaa accaacagta ttggtgagga 27420
 cagaaacacc acggtaacga agaatacgg cttttccgta aaactggcgc agacgatcaa 27480
 tatcggcacc acttatcggt tagatgttg cgatcaattc acgttcgct gcggcaatgc 27540
 ggcgttggt ttacataagg acggtccat tgagttttgt ggcaagcaac tgatgttaca 27600
 taccagcgat gtcatgcaac tgattggtaa aggtattgat atgaaccgg atggcggcac 27660
 agccgtaacc gccgatgata ttgccccct tctcacctct gattgatctg aattaaacct 27720
 ggagttctca tggatcgacc ataccgata caggaagggg gttttgtcct gcctgaaaca 27780
 ttacggatc gcagcgtaa ttttttatc ctggagggca atgaacgaac atcgcccagc 27840
 ctgaatattt cccgcgatac gctaaaacct gatgaagacc tgccgccta tattgaccgc 27900
 cagattgcac tgatgaaaaa aaatctcggc cagcaccggg tattgtcgcg agcgctgca 27960
 caggcaggaa cgggcaatga tgcccttatg ggggaacaaa ttgccgccac ccataaatcc 28020

gggaaaacgg aagtgtacc cgtcaggcc gggtttattg caaccctgg aaggtactg 28080
gtcttcaccc tgaccagtcc cgtcctttt gatgataaag cagacctact ctggaacacc 28140
tggctggcag gctttcagcc ggataaaaac gaataatcac acggaggtgt gaccatgtat 28200
gaagcagccc gtgtggatga tcctatctac cacaccagcg cgctcgccgg gtttcttata 28260
ggcgctatca tcggcatcgc cattatcgcg cttgccgcct ttgccttctt tagctgcggt 28320
tttcttgccg ggctgattct gggttttatg gccgatcaaa tagcctccgg ggtattgcaa 28380
ctgggcgagg ccatacggcg ctcacccac cacacggcag gaaaaatcct caccggttcg 28440
gagaatgtca gcaccaacag tcgcccggcg gcgcgcgcgg tactgagtag ggtgaaatgc 28500
gataaccata tcgcagaaaa acgcatcgcc caagggtcgg aaaatatcta catcaacagc 28560
cagccccggc cccgtaagga tgaccacacc gaatgcgacg cggtgattga agacggttcg 28620
ccgaatgtgt ttctcggcg cggcacacag acggtactgg aaatcagttc tgaaattccg 28680
gactggctgc gcaagggtgt ggatgtattg tttgtcgtgg cgagtctgct cggcgggctg 28740
gccggggcgt ggcggcaggc ggcaaagctg gggacgaaat ttggcactaa atgtgccgct 28800
aagtttatcg gcggggagct tgcgggatg gccgtgggtg aggtatcag cgggctgttc 28860
agcaatccgg tggatgtgac caccgggcag aaaatcctgc tgccggaaac ggacttcacc 28920
ctgcccggtc gcctgccggt cacctgctcg cgtttttacg ccagccacct ggaaactgtg 28980
ggactgttgg gacggggctg gcggctgaac tgggaaacca gcctgcgcga tgacgatgaa 29040
cacatcacgc tgaccggcgt acaggggcgg gaactgcgtt acccgaaaac gatgctgacg 29100
cccggccacc agatatttga cccggaagaa cagttatacc tcagccgctt gcatgacggg 29160
cgttacgtgc tgcattacac cgatcgcagc tattacgtat ttggtgattt tgacagtgc 29220
ggcatggcat acctgctgtt tatggagacg ccgcaccgcc agcgcatgtt cttcgggcac 29280
gaaggaggca gactggtacg gatagcctcc agcagcgggc atcacctgtt actgcaccgc 29340
acacagaccc cggcagggga gcggctgtcg cgaattgaac tggcgcaggc cggcaccgct 29400
ggcaatctgg tggagtaccg gtatgacgat aacggtcaac tgaccggcgt ggtgaaccgg 29460
gcgggaacgc aggtgcgtca gtttgcttat gaaaacgggc tgatgacggc gcacagcaat 29520
gcgacggggt tcacctgccg ctaccgctgg caggaaactcg acggcgccgc gcgcgtgacg 29580
gagcagcaca ccagtgcgg cgaacattac cgctttgact atgattttgc cgcaggcacc 29640
accaccgtca ccggcaggca gggggagaca tggcagtggt ggtacgacag ggaaacgtat 29700
atcaccgcgc accggacgcc gggcggtgga atgtaccgct tcacgtacaa cgaagaccac 29760
ttccctgtca acattgagct gccggcggt cgcacgggtg cgtatgaata tgacatccag 29820
aaccgggtgg tgaagacgac agatccggaa ggccgggtga cgcagacgca gtggaacggc 29880

gagttcgacg aaatcacgcg cacggcgctg gacgatgacg ctgtctggaa aacgcagtac 29940
aacgcccacg gccagccagt gcaggagacg gacccggaag ggcgggtgac gcagtagcgt 30000
tacgatgaac aggggagat gtgcagccgg acggatgcgg cgggcggcac gcagggcggc 30060
gtgcggcggg agacgcagca gcgggatgcg ctgggcccgc tggttacggac ggagaatgaa 30120
cacggccagc ggacgttcag ctacaaccgg ctggaccaga taacggcagt gacgctcacg 30180
cccacggagg cggggcaaca gcagcaccgg atgcaggccg acacgggtgcg ttttgagtat 30240
gaccgcagcg gctggctgac ggcggagcac gcggggaacg gtagcatatg ttatcagcgc 30300
gacgcgctgg gcaaccgcgac ggacatcacg ctgccggacg ggcagcacct gacgcatctg 30360
tattacggga gcgggcatct gttacagacg gcgctggacg gcctgacggg gagcgagtat 30420
gagcgcgaca gcctgcaccg tcagataatg cgcacgcagg ggcagcttgc gacgtacagc 30480
ggctatgacg acgacgggt gctgagctgg cagcgcagtc tggcgccggc cagtggccct 30540
gttcttctcg gccagcgcgc ggcgcggcag ggctgcgtga cgtcgaggga ctattactgg 30600
aacaaccacg gcgaggtggg cacgattgac gacggcctgc gtggcagcgt ggtgtacagc 30660
tatgacagaa gcggttacct gaccgggcgc tcaggtcaga tgtatgacca tgaccgttat 30720
tattacgata aggcggggcaa cctgctggat aacgaagggc agggagcggg gatgagcaac 30780
cggctgccgg gctgtggtcg tgaccgttac ggctataacg agtggggcga gctgaccacg 30840
cggcgcgacc agcaactgga gtggaacgcg caggggcagc tgacgcgggt catcagcggc 30900
aacacggaga cgcactacgg ctacgatgcg ctggggaggc gaaccgcga ggacgacgtac 30960
gggcggcaca cgggccatac ggcgcggagc cggacggact ttgtgtggga ggggttcagg 31020
ctgttgacg agaacgtgca gcagcagggc tggcggacct atctgtacga tgcggaacag 31080
ccgtacacgc cgggtggcgag cgtgacgggg cggggagaaa gcaggcaggt gtggtattac 31140
cacacggatg tgacgggcac gccgcaggag gtgacggcgg cggacggaac gctggtgtgg 31200
gcgggggtata tcaggggggt tggagagaat gcggcggaca tcagcaacag cggggcgtag 31260
tttcaccagc cgctgcggct gccggggcag tttttgacg acgagacagg gctgcattac 31320
aatctgttca gatattatgc accggagtgt ggacggtttg tcagtcagga tccgatcggg 31380
ctgaggggcg ggttaaactt ttatcagtat gcgccaaatc ctctcaaata tatagacca 31440
cttggtttta cgcgcactgt tgggcgatgg atggggcctg cggaatatca gcaaagtctt 31500
gatactggga cagtagtaca aagttcaaca gggacaactc atgttgcta cctgctgat 31560
atagatgctt ttggtaagca agcaaaaaat ggtgctatgt atgttgaatt tgatgtgcct 31620
gaaaaatcat tagtacctac aaatgaagga tgggcaaaaa tagtagggcc agattctatc 31680

gaagggcgat tagctaaacg aaagggttg cctgttcctg aaatgccaac agcagaaaac 31740
 ataactgtaa ggggcgagaa aattaatggg gaagttgaag caaaatgcta aataaattta 31800
 aattgtgggt gagcaaacat actgattata cggtaattca taatgaaaat gatttatctt 31860
 acagtattat tatagatttt gaagatgacc ggtatatatc aagatttact gtatgggatg 31920
 acctaagctg tatgtcagaa gtaatggatg tggatactgg tttatataaa ttaaacaaga 31980
 gaaacgaatt ttctacattt gatgaacttc tggatatatt tgatgatttt atgataagta 32040
 ttaaataata gttggccggg taagaagtta actcttcccg gctgttttat tatctaacc 32100
 ccatcaatcc ggagacgcgc taccggtagc atgcgctggg caggcgggtg agcaaggcga 32160
 cgtacgggcg gcacacgggc catacggcgc ggagccggac ggactttgtg tgggaggggt 32220
 tcaggctgtt gcaggagaac gtgcagcagc agggctggcg gacatatctg tacgatgcgg 32280
 aacagccgta cacgccggtg gcgagcgtga cgggaaaggg agaaagcagg cagggtgtgg 32340
 attaccacac ggacgtgacg ggcacgccgc aggaggtgac ggcgccggac ggaacgctgg 32400
 tgtgggcggg gtatatcagg ggggttgagg aaaatgcggc ggacatcagc aacagcgggg 32460
 cgtactttca ccagccgctg cggctgccgg ggcagtattt tgacgacgag acagggctgc 32520
 attacaatct gttcagatat tatgcaccgg agtgtggacg gttcgtcagt caggatccga 32580
 ttgggctggc ggggggggct gaatctttac cagtatgcgc ctaatccgat tagatggatc 32640
 gatcctttag gacttgctat cctggagcat caatctaatt ttgatgcggc aaggagaacc 32700
 ggatttgaaa atgcgggtat gacaaacct gaggatgtca ctttctcgaa agtcgatccc 32760
 aaaactggta ctgttggtga gtttaaaggt ccaaagggg ctaaagttgc ttatgatgca 32820
 cctcatgcag atatggatgt gacagcaggg catgataaac cacatggttg ttggcaatcc 32880
 gcaggaaaaa gaggttccgg aggagctaata agaggtaata ttacttatga tggcccacaa 32940
 catccgcacg gctctgactc taaggagat gataaatgtt aaattcaaata atgtctgaac 33000
 ttagaatcga actggagaat gcgattaaaa atctcggtat tcatgattat cgtgtcgata 33060
 aaccgaaca aatcgtttct gagataaaag agatatatgt taatggtaat cctagaacct 33120
 ggtgggttatc attaaaacat agacaatatg tcttttctta taccgataat tctggatata 33180
 aaaacatatc acaaatagta agtaaacaac tcaatgaaag caatgtaatc acaaacata 33240
 tatttttgat tgctgatgaa gataatgagc aaatatatgt atataacgtt cctcttaact 33300
 ccctgcctga aattatagaa aattgcagat attttgaata ttatggtgca gatcatgaac 33360
 tatcttggtc tatatgtgaa aatgatcatg gtgatttgat tgtatgctca accattaagt 33420
 aaagcgcgag tgctctttag cgatatagtt gcccatattt aggcgttact agccgaagat 33480
 ggcgcgattg tctggcaggg gaaacagcaa ttctgaggtc aggaagatag cataacccat 33540

taaccgggat agatccgcta gacctgaatc cagttgatgc gacaggttat aggggtttatg 33600
 gttattttgct cctggagcaa ataaacctta ttacattggt attactaatg atatggtttg 33660
 acgaagggcc gagcattaaa gcaactggcag gttatcaaaa gaaaatggaa ggatgctgcc 33720
 atttgatgaa aatgtaatct aatggaaagt cagaggttac gagaaatatt atatagagaa 33780
 atataaaacc agaaccggaa ccatagggtga aaaaattccc tcaacaaata gagagaataa 33840
 atataattca tttgatcatg ggcgaacaga tcccagcgca caagcattta aagactctta 33900
 aaatagtaag ggagttgggt cgggtggagg aaaatgcgga tgagtgatta agaatttttg 33960
 ggctgtgata agaagtcgag aacaatgctg cgttttgtga agcccgagga catattttgt 34020
 tttaaattag atgaagatag atattgtttt gggcgaatta taacactaat gactgtcgga 34080
 catctttctg aattatttga tataattaa aaacccctg gaataacaga gttagaaatt 34140
 agtaatgcaa ggcgaattat tgaaccaatt atagtggata catattcttt atttgataag 34200
 aaattagaaa atggaagtga ctggagaatt attggtcatc aggttaatta caatccaaaa 34260
 aatttagatg gtatctattt tgcacttga ataggtgatt cctgtaaaaa gaaagactgt 34320
 tacggaaatg attttctcat ttcagaaagt gagtggaaaa cacttcctaa attatctcct 34380
 aaaggggggt ttgatatcaa aaaacggctt gaaattgcct gaaaatgaaa ataaaaagcc 34440
 gggaaagatc ttttgtcttc ccgattttta ttatttaatc cccgttcacc acattattta 34500
 cccccgctt aatatgcttc atcgactttt tcacctgata aagctccttc cgtagatccc 34560
 tcacttcgtc cgtctctgca atcaggatca aacacccctc ggagatcttc acggtgacgc 34620
 cgtgcccgggt ctcaaattccc gcttcttctt gccagtcacc cttaagggtc tggctgggga 34680
 tttgtgagta acaggcggtc atgcagggtt cgctgttgat atggcggacg ctgacgccgg 34740
 aggcattcat atttgctgac taaataaatt cttatttatc cgccggatgc tggctgattg 34800
 tggagctcag ggtgagtga tatgggcgag acatcctggc accgcgcctc cctctccccg 34860
 gccagccccg gccggtgatg agcagccggc tgccggggcc gtattttgac gatgaaacgg 34920
 gctgcatta aaatctgttc agatattatg taccggagtg tggctgggtc gtcagtcagg 34980
 atccaatagg gctgaaaggg ggatggaacc gatatcattc tccgctgaat cctattacag 35040
 atagtgatcc tcttggcctt attacttggt gtgctgatag aggtgattct ggcaagtat 35100
 taagatgaaa aatggttgaa aataattgag ttagtgaatg ctcatztatg cctttaccta 35160
 caaaaagtaa cggatttgct tgctggaatt gtgttaacga atgcaaataa aatcaatgct 35220
 gttgatgttt ttactacagg aataagtatg gttaacgata aggatacagc tatattaatt 35280
 agtaatttaa tgttgagatt cggtaaggag cttgatgaat ctgttgctgt tgttcagtc 35340

cgttgtgatg aggatgaatt tcatgtatat cgagaaacgg ttgggttttat catgggtgaa 35400
atgcttatta aaataatgaa tccattatat gaaaaacatc cagaaataaa accaaaagga 35460
ttgaaacaaa acatctggaa ccggatgaat aatgtgtaaa agccggaggg gttatctttt 35520
cccggctttt tattatcaat tactcattaa ctctgttcc gttcttttgc gtttaatcac 35580
cggaatatct ccggtattgt tcagcgcccc ggaaatgttt ttaaccactg ttctgcactc 35640
cgtttattaa tgcggggttac gcccatccct tcaatacagc caaagagtcc gtgggtatgc 35700
tgcgcggtga tcacgatgca atccctcatt acccgcacct tgacgggcat ccggttaata 35760
aaccocgcct gcggcatcca ttcaccgtac aaacgaactg aaagtctctc aacgcgtgag 35820
tatgtaagta tcccgcataa tcgagccatt cacatttaga gatcatccga cataatcaat 35880
ctgccaacgc aggagatcgc tatgcgtaaa gcccgattta ctgcgcacca gatcatcgct 35940
gtgattagat cagtcgaatc cggacggact gttaaagatg tctaccggga gcccggtatt 36000
tctgaagcca ccagggacaa ctggaggtct ggatacggcg gcagggatac gcgtggaatc 36060
acaaaaggct ccaccgtatt tactgtctgc tcaagctgaa ttttcgccgt aagggtaaac 36120
aacggctgcc ggtacgcaat cctcgcacc tggtcacgcc ggaagcgctg aaccagagct 36180
ggctctgtgg cgctcgtttt gcacgttcaa tgttgttgat gactgtaatc gtgaagcgtt 36240
gtcgattgaa atcgatctga atctgccagc tctgcgagtg gtccgtgtac tcgacaggat 36300
tacagcaacc gcggttatct ggccatgctg cgtatggata agggaccgga atttatctcg 36360
ctggcactgg ctgaatgggc aaagaaacat gcagtaaagc tggcgtttat ccagccgggt 36420
aagccgaaga aaaacgtttt catcacgcgc ttaaccgga cataccgtac agaaatactc 36480
aattcttata tgttcagaac gctgaatgag gtgtgggaaa ttacggataa agggttatca 36540
gaatataact gcgaacgtcc acatgaatcg cggaacaata tgataccgaa ggaataccgc 36600
caataacggt atctggccgg aatcttaaaa atgcatggaa ctaaaacggg tctatttaca 36660
ggggcacctg cgatgaattt cgctgcactg aaaagcgata ccggatgaga gctgcttcaa 36720
attaatgtgc catgttcacg gggaggttgt gcgacgtttg cataatccag caagaactga 36780
aaggaagggg agagcttttt catgcctgta taatcagtct ggctgtgtc agtcagctct 36840
tagtgttgag actctcgttg gagcgttata attgcttttc tgtttcggaa aacaagattt 36900
tccattaaag atcttccctg cgaggaaaag ttaactaata atcttaccgt cgagtttaga 36960
gatgtatgtt taaatataaa caatgttgca acgatgcctg ataattatcc tctcttcgaa 37020
gataagtttc ccacaccag tgtagtaggt gtcattgtaa tgttatcact tgaatgtaaa 37080
tggaaggat aattgctttt tgactggcat tctattccac cctgacaaca cgatgttaac 37140
atcaacactg tttatattgg caataacgca atttttttca gattaagagg tgctctgata 37200

tca gta tat gcc gtc aac cag caa tta aat tca gcc act aaa tta ttc	38059
Ser Val Tyr Ala Val Asn Gln Gln Leu Asn Ser Ala Thr Lys Leu Phe	
190 195 200	
agc gtg aag ctg ggg gct aca cga gtg att tat cac gct ggt acg gct	38107
Ser Val Lys Leu Gly Ala Thr Arg Val Ile Tyr His Ala Gly Thr Ala	
205 210 215	
gga gcg acg ctc tcg gtg agc aac ccg cag aat tac cct att ttg gtt	38155
Gly Ala Thr Leu Ser Val Ser Asn Pro Gln Asn Tyr Pro Ile Leu Val	
220 225 230 235	
cag tct tca gtc aaa gca gca gac aaa agt tcg cct gct ccc ttt ttg	38203
Gln Ser Ser Val Lys Ala Ala Asp Lys Ser Ser Pro Ala Pro Phe Leu	
240 245 250	
gtg atg ccg cct cta ttt cgt tta gaa gca aac cag cag agt caa ctg	38251
Val Met Pro Pro Leu Phe Arg Leu Glu Ala Asn Gln Gln Ser Gln Leu	
255 260 265	
cgt att gtc cgt act ggt ggt gac atg cca acg gat cgt gag act tta	38299
Arg Ile Val Arg Thr Gly Gly Asp Met Pro Thr Asp Arg Glu Thr Leu	
270 275 280	
cag tgg gtc tgt ata aag gcg gta cca ccc gaa aat gaa ccg tcg gat	38347
Gln Trp Val Cys Ile Lys Ala Val Pro Pro Glu Asn Glu Pro Ser Asp	
285 290 295	
aca cag gct aag ggc gcg acc ctt gac ctc aat ttg tcc atc aac gcc	38395
Thr Gln Ala Lys Gly Ala Thr Leu Asp Leu Asn Leu Ser Ile Asn Ala	
300 305 310 315	
tgt gat aag ctg att ttc cgc ccg gat gcc gtg aag ggg acg ccg gaa	38443
Cys Asp Lys Leu Ile Phe Arg Pro Asp Ala Val Lys Gly Thr Pro Glu	
320 325 330	
gat gtt gca gga aat tta aga tgg gtg gag acg ggc aac aaa ctt aag	38491
Asp Val Ala Gly Asn Leu Arg Trp Val Glu Thr Gly Asn Lys Leu Lys	
335 340 345	
gtg gag aac ccc acc ccg ttt tac atg aat tta gcc tct gtc aca gta	38539
Val Glu Asn Pro Thr Pro Phe Tyr Met Asn Leu Ala Ser Val Thr Val	
350 355 360	
ggg gga aag ccc att aca ggg ctt gag tat gtc ccc ccc ttt gct gac	38587
Gly Gly Lys Pro Ile Thr Gly Leu Glu Tyr Val Pro Pro Phe Ala Asp	
365 370 375	
aaa aca cta aat atg cca ggt agt gcc cat ggt gat atc gag tgg aga	38635
Lys Thr Leu Asn Met Pro Gly Ser Ala His Gly Asp Ile Glu Trp Arg	
380 385 390 395	
gtt att aca gac ttt ggt ggt gaa agt cat ccg ttc cac tac gtt ctt	38683
Val Ile Thr Asp Phe Gly Gly Glu Ser His Pro Phe His Tyr Val Leu	
400 405 410	
aaa taa atccaggggc ttagcggcag aaa atg aag ttc aaa caa cct gcc ttg	38736
Lys Met Lys Phe Lys Gln Pro Ala Leu	
415 420	
cta ctg ttc atc gcg gga gtg gtt cat tgc gca aat gcg cac act tac	38784

Leu	Leu	Phe	Ile	Ala	Gly	Val	Val	His	Cys	Ala	Asn	Ala	His	Thr	Tyr		
			425					430					435				
aca	ttc	gat	gca	tca	atg	ttg	ggc	gat	gca	gcg	aaa	ggg	gtt	gat	atg	38832	
Thr	Phe	Asp	Ala	Ser	Met	Leu	Gly	Asp	Ala	Ala	Lys	Gly	Val	Asp	Met		
		440					445					450					
tcg	ctc	ttt	aac	cag	ggg	tta	caa	cag	cca	ggg	act	tat	cgc	gtg	gac	38880	
Ser	Leu	Phe	Asn	Gln	Gly	Leu	Gln	Gln	Pro	Gly	Thr	Tyr	Arg	Val	Asp		
		455					460				465						
gtg	atg	gtg	aat	ggg	aaa	cgt	gtc	gac	acc	cgt	gat	gtg	gtg	ttc	aaa	38928	
Val	Met	Val	Asn	Gly	Lys	Arg	Val	Asp	Thr	Arg	Asp	Val	Val	Phe	Lys		
470					475					480					485		
ttg	gaa	aag	gat	ggg	caa	gga	acg	cct	gtt	ctg	gct	cct	tgt	ttg	acg	38976	
Leu	Glu	Lys	Asp	Gly	Gln	Gly	Thr	Pro	Val	Leu	Ala	Pro	Cys	Leu	Thr		
				490				495						500			
gtc	agt	cag	ctt	tca	cgc	tac	ggc	gta	aaa	acg	gaa	gat	tac	cct	cag	39024	
Val	Ser	Gln	Leu	Ser	Arg	Tyr	Gly	Val	Lys	Thr	Glu	Asp	Tyr	Pro	Gln		
			505					510					515				
ttg	tgg	aaa	gca	gca	aag	ccc	cca	gat	gag	tgt	gcg	gat	ctg	acc	gcc	39072	
Leu	Trp	Lys	Ala	Ala	Lys	Pro	Pro	Asp	Glu	Cys	Ala	Asp	Leu	Thr	Ala		
		520					525					530					
att	cca	cag	gct	aaa	gcg	gta	ctg	gat	atc	aat	aat	cag	caa	ctg	caa	39120	
Ile	Pro	Gln	Ala	Lys	Ala	Val	Leu	Asp	Ile	Asn	Asn	Gln	Gln	Leu	Gln		
		535					540					545					
ctg	agt	att	ccg	cag	ttg	gcg	ttg	cgt	ccg	gaa	ttt	aag	ggg	atc	gct	39168	
Leu	Ser	Ile	Pro	Gln	Leu	Ala	Leu	Arg	Pro	Glu	Phe	Lys	Gly	Ile	Ala		
550					555					560					565		
cca	gaa	gat	ctt	tgg	gat	gat	ggg	att	ccg	gcg	ttt	ctg	atg	aac	tac	39216	
Pro	Glu	Asp	Leu	Trp	Asp	Asp	Gly	Ile	Pro	Ala	Phe	Leu	Met	Asn	Tyr		
				570				575						580			
agt	gcg	agg	aca	acg	cag	acg	gat	tac	aaa	atg	gat	atg	gtg	ggg	cgt	39264	
Ser	Ala	Arg	Thr	Thr	Gln	Thr	Asp	Tyr	Lys	Met	Asp	Met	Val	Gly	Arg		
			585					590						595			
gac	aac	tct	tcc	tgg	gta	caa	ctg	caa	ccg	gga	atc	aat	ata	ggg	gcg	39312	
Asp	Asn	Ser	Ser	Trp	Val	Gln	Leu	Gln	Pro	Gly	Ile	Asn	Ile	Gly	Ala		
		600					605					610					
tgg	cgt	gtc	cgc	aat	gcg	acc	agc	tgg	cag	cgg	agt	agt	caa	ctg	tcg	39360	
Trp	Arg	Val	Arg	Asn	Ala	Thr	Ser	Trp	Gln	Arg	Ser	Ser	Gln	Leu	Ser		
		615				620						625					
ggg	aag	tgg	cag	gca	gca	tat	acc	tat	gct	gag	cgt	gga	ctg	tac	tca	39408	
Gly	Lys	Trp	Gln	Ala	Ala	Tyr	Thr	Tyr	Ala	Glu	Arg	Gly	Leu	Tyr	Ser		
630					635					640					645		
cta	aaa	agt	cgt	ctg	act	ctg	ggg	caa	aag	act	tcg	cag	ggg	gag	ata	39456	
Leu	Lys	Ser	Arg	Leu	Thr	Leu	Gly	Gln	Lys	Thr	Ser	Gln	Gly	Glu	Ile		
				650				655						660			
ttt	gat	agt	gtg	cca	ttt	acc	ggg	gtg	atg	ttg	gca	tcg	gat	gac	aac	39504	
Phe	Asp	Ser	Val	Pro	Phe	Thr	Gly	Val	Met	Leu	Ala	Ser	Asp	Asp	Asn		

665						670						675						
atg	gtg	ccc	tac	agt	gag	cgg	cag	ttt	gct	ccg	gta	gtg	cgt	ggg	att	39552		
Met	Val	Pro	Tyr	Ser	Glu	Arg	Gln	Phe	Ala	Pro	Val	Val	Arg	Gly	Ile			
		680					685					690						
gcc	cgc	acg	cag	gct	cgg	gtg	gag	gtc	aaa	cag	aat	ggg	tac	acc	att	39600		
Ala	Arg	Thr	Gln	Ala	Arg	Val	Glu	Val	Lys	Gln	Asn	Gly	Tyr	Thr	Ile			
		695				700					705							
tac	aac	acc	act	gtg	gcg	ccc	gga	ccg	ttt	gca	ctg	cgg	gat	ctg	tcg	39648		
Tyr	Asn	Thr	Thr	Val	Ala	Pro	Gly	Pro	Phe	Ala	Leu	Arg	Asp	Leu	Ser			
710					715					720					725			
gta	aca	gac	agt	agt	ggg	gat	ctg	cat	gtc	acc	gtg	tgg	gag	gcc	gat	39696		
Val	Thr	Asp	Ser	Ser	Gly	Asp	Leu	His	Val	Thr	Val	Trp	Glu	Ala	Asp			
				730					735					740				
ggc	agt	aca	caa	atg	ttt	gtg	gtg	ccg	tat	cag	acc	ccg	gcg	ata	gca	39744		
Gly	Ser	Thr	Gln	Met	Phe	Val	Val	Pro	Tyr	Gln	Thr	Pro	Ala	Ile	Ala			
			745					750					755					
ctg	cac	cag	gga	tat	ttg	aag	tac	agc	ctg	ttg	gcg	ggc	cga	tac	cga	39792		
Leu	His	Gln	Gly	Tyr	Leu	Lys	Tyr	Ser	Leu	Leu	Ala	Gly	Arg	Tyr	Arg			
		760					765					770						
tcg	tca	gac	tct	gca	acg	gat	aag	cgg	cag	atc	gcg	cag	gct	acg	ttg	39840		
Ser	Ser	Asp	Ser	Ala	Thr	Asp	Lys	Arg	Gln	Ile	Ala	Gln	Ala	Thr	Leu			
		775				780					785							
atg	tat	ggg	ctg	ccg	tgg	aat	ctc	act	gca	tac	ggc	ggg	ata	cag	agt	39888		
Met	Tyr	Gly	Leu	Pro	Trp	Asn	Leu	Thr	Ala	Tyr	Gly	Gly	Ile	Gln	Ser			
790					795				800					805				
gca	acg	cat	aat	caa	gct	gca	ttg	ctt	ggg	ttg	ggg	gga	tct	ctc	ggg	39936		
Ala	Thr	His	Asn	Gln	Ala	Ala	Leu	Leu	Gly	Leu	Gly	Gly	Ser	Leu	Gly			
				810					815					820				
cgg	tgg	ggg	agt	tta	tct	gtc	gat	gga	agc	gac	aca	cac	agt	cag	cgt	39984		
Arg	Trp	Gly	Ser	Leu	Ser	Val	Asp	Gly	Ser	Asp	Thr	His	Ser	Gln	Arg			
			825					830						835				
cag	ggg	gag	gcg	gta	cag	caa	gga	gcc	tcc	tgg	cga	ctg	cgt	tac	agc	40032		
Gln	Gly	Glu	Ala	Val	Gln	Gln	Gly	Ala	Ser	Trp	Arg	Leu	Arg	Tyr	Ser			
		840					845					850						
aac	cag	ctg	act	gcg	acg	ggg	aca	aat	ttt	ttt	ctg	acg	aga	tgg	cag	40080		
Asn	Gln	Leu	Thr	Ala	Thr	Gly	Thr	Asn	Phe	Phe	Leu	Thr	Arg	Trp	Gln			
		855				860					865							
tat	gcc	tcg	cag	ggc	tat	aac	acc	cta	tcc	gat	gtg	ctc	gac	agt	tat	40128		
Tyr	Ala	Ser	Gln	Gly	Tyr	Asn	Thr	Leu	Ser	Asp	Val	Leu	Asp	Ser	Tyr			
870					875					880					885			
cga	cat	aat	ggc	aac	cgt	cta	tgg	tcg	tgg	cgg	gaa	aat	ttg	cag	ccg	40176		
Arg	His	Asn	Gly	Asn	Arg	Leu	Trp	Ser	Trp	Arg	Glu	Asn	Leu	Gln	Pro			
				890					895					900				
agc	tcg	cgt	act	acc	ctg	atg	ttg	agt	cag	tca	tgg	ggg	agg	cat	ttg	40224		
Ser	Ser	Arg	Thr	Thr	Leu	Met	Leu	Ser	Gln	Ser	Trp	Gly	Arg	His	Leu			
			905				910						915					

ggc aat ctg agt tta acc ggt tcc cgt acc gac tgg cgt aat cgc ccc Gly Asn Leu Ser Leu Thr Gly Ser Arg Thr Asp Trp Arg Asn Arg Pro 920 925 930	40272
ggt cat gat gac agc tac gga ctg agt tgg gga acc tct atc gga ggg Gly His Asp Asp Ser Tyr Gly Leu Ser Trp Gly Thr Ser Ile Gly Gly 935 940 945	40320
ggc tcg ctg tca ttg aac tgg aat caa aac aga acg ctg tgg cgc aat Gly Ser Leu Ser Leu Asn Trp Asn Gln Asn Arg Thr Leu Trp Arg Asn 950 955 960 965	40368
ggc gcg cac cgt aaa gag aac ata acc agc ctg tgg ttc agt atg cca Gly Ala His Arg Lys Glu Asn Ile Thr Ser Leu Trp Phe Ser Met Pro 970 975 980	40416
tta agc cgc tgg acg ggg aat aat gta agt gct agt tgg cag atg act Leu Ser Arg Trp Thr Gly Asn Asn Val Ser Ala Ser Trp Gln Met Thr 985 990 995	40464
tca cca tca cac ggt ggt cag acg caa caa gtg ggg gtc aac gga gag Ser Pro Ser His Gly Gly Gln Thr Gln Gln Val Gly Val Asn Gly Glu 1000 1005 1010	40512
gca ttc agt cag caa ctg gat tgg gag gtg cgt cag agt tac cgt gcc Ala Phe Ser Gln Gln Leu Asp Trp Glu Val Arg Gln Ser Tyr Arg Ala 1015 1020 1025	40560
gat gcc ccg cca ggt ggt ggt aat aac agc gca ttg cac ttg gca tgg Asp Ala Pro Pro Gly Gly Gly Asn Asn Ser Ala Leu His Leu Ala Trp 1030 1035 1040 1045	40608
aat ggg gat tac ggc ctg tta ggt ggt gac tat agc tac agc cgg gcg Asn Gly Asp Tyr Gly Leu Leu Gly Gly Asp Tyr Ser Tyr Ser Arg Ala 1050 1055 1060	40656
atg cgc cag atg gga gtc aat atc gcg gga ggt ata gtt atc cac cat Met Arg Gln Met Gly Val Asn Ile Ala Gly Gly Ile Val Ile His His 1065 1070 1075	40704
cat ggt gtg acg ctg ggg caa cct ttg caa ggc tca gtg gcg ctg gtt His Gly Val Thr Leu Gly Gln Pro Leu Gln Gly Ser Val Ala Leu Val 1080 1085 1090	40752
gaa gcg cca ggg gcc tcg ggg gtg cca gtt ggc ggc tgg cct ggc gtt Glu Ala Pro Gly Ala Ser Gly Val Pro Val Gly Gly Trp Pro Gly Val 1095 1100 1105	40800
aag acg gat ttt cgt ggc gac acc aca gtg ggc aac ctg aac gtc tat Lys Thr Asp Phe Arg Gly Asp Thr Thr Val Gly Asn Leu Asn Val Tyr 1110 1115 1120 1125	40848
cag gag aat aca gtc agc ctc gat ccg tcg cga cta ccg gat gac gca Gln Glu Asn Thr Val Ser Leu Asp Pro Ser Arg Leu Pro Asp Asp Ala 1130 1135 1140	40896
gag gtc aca caa acc gat gtg cgc gtg gtg cca acc gaa ggg gcg gtg Glu Val Thr Gln Thr Asp Val Arg Val Val Pro Thr Glu Gly Ala Val 1145 1150 1155	40944

gtg gaa gcg aag ttt cac act cgc atc ggg gcc agg gca ctg atg acg Val Glu Ala Lys Phe His Thr Arg Ile Gly Ala Arg Ala Leu Met Thr 1160 1165 1170	40992
ctg aaa cgg gaa gat ggt agc gcc att cct ttc ggg gcg cag gtt aca Leu Lys Arg Glu Asp Gly Ser Ala Ile Pro Phe Gly Ala Gln Val Thr 1175 1180 1185	41040
gtc aat ggg cag gat ggc agt gct gct ctg gtg gat act gat agc cag Val Asn Gly Gln Asp Gly Ser Ala Ala Leu Val Asp Thr Asp Ser Gln 1190 1195 1200 1205	41088
gtt tat ctc act ggt ttg gcg gat aag ggc gaa ctg acg gtg aaa tgg Val Tyr Leu Thr Gly Leu Ala Asp Lys Gly Glu Leu Thr Val Lys Trp 1210 1215 1220	41136
gga gca cag caa tgt cgg gtt aac tac cgc cta cct gcc cac aag gga Gly Ala Gln Gln Cys Arg Val Asn Tyr Arg Leu Pro Ala His Lys Gly 1225 1230 1235	41184
atc gcg ggc ttg tat caa atg agc ggt ctc tgc aga tag ccgattctga Ile Ala Gly Leu Tyr Gln Met Ser Gly Leu Cys Arg 1240 1245 1250	41233
aggagagaat a atg tgg atg aaa ata cag cga gtg aaa acg gtt atc tat Met Trp Met Lys Ile Gln Arg Val Lys Thr Val Ile Tyr 1255 1260	41283
agc gta agc tta ctg gtc gct gcc agt agc ttg gtg ccg ata gcg aac Ser Val Ser Leu Leu Val Ala Ala Ser Ser Leu Val Pro Ile Ala Asn 1265 1270 1275	41331
gcc gca gaa aaa ctt cag aca acg cta cgt gta ggt act tac ttt cgg Ala Ala Glu Lys Leu Gln Thr Thr Leu Arg Val Gly Thr Tyr Phe Arg 1280 1285 1290 1295	41379
gct ggg cac gtg cca gat ggg atg gtg ctt gcg caa ggc tgg gtg act Ala Gly His Val Pro Asp Gly Met Val Leu Ala Gln Gly Trp Val Thr 1300 1305 1310	41427
tat cac ggc agt cac agc ggg ttt cgg gta tgg agc gat gag caa aag Tyr His Gly Ser His Ser Gly Phe Arg Val Trp Ser Asp Glu Gln Lys 1315 1320 1325	41475
gcg ggt aac acg cct acc gta ttg ctg ctg agc ggg caa cag gat cct Ala Gly Asn Thr Pro Thr Val Leu Leu Leu Ser Gly Gln Gln Asp Pro 1330 1335 1340	41523
cgc cat cac att cag gtt cgc ctg gag ggc gag ggg tgg caa cca gat Arg His His Ile Gln Val Arg Leu Glu Gly Glu Gly Trp Gln Pro Asp 1345 1350 1355	41571
acg gtg agt ggt cgt ggc gcc att tta aga acc gct gca gat aac gcc Thr Val Ser Gly Arg Gly Ala Ile Leu Arg Thr Ala Ala Asp Asn Ala 1360 1365 1370 1375	41619
agt ttc agt gtg gtc gtt gat ggc aat cag gaa gtg cct gcg gac acc Ser Phe Ser Val Val Val Asp Gly Asn Gln Glu Val Pro Ala Asp Thr 1380 1385 1390	41667
tgg acg ctg gat ttt aag gcc tgt gca ttg gcg cag gag gat acg tag	41715

Trp Thr Leu Asp Phe Lys Ala Cys Ala Leu Ala Gln Glu Asp Thr
 1395 1400 1405

ccgtctgttc cactcatact tcctgtatca gtgaataagg cgaatactgt atatcagggt 41775
 cgaggatgta ttgaaaattt actcaaggca aatatgggta gcaatcttca ttgcagaaaa 41835
 atcaattgct gattgatttt aaaatctgaa attcatcttt tttgtaggga gaggaatatt 41895
 atgtttcgta aaataattta acggacactg atgcctttat tatttctgtt ttgtggattt 41955
 gtttgattca ctcttttgtg agcggggctt tataagcccc tggataaagg tttgtttatc 42015
 tgtcagggtat tcatatgtga tatttaagat tttatgttta gggggctttt tcttgtcccc 42075
 tcaggcgtaa aaataattta ttttttacat aaaggaataa agcatatgtc ttatgcacga 42135
 catttaccgg tattaatgta tcaccatgtc agtgataaac ccggacagat aaccttatct 42195
 ccccgtagct tccgggcgca gatgaaatgg ctggccgaat ctggctggaa aaccgttact 42255
 gcctgcagag gtggaagctt tttatcatgg tgcaagattg cctcgtaaaa gcgtcatgct 42315
 gacctttgat ggcggctggc tggataactg gttgcagggt tttccgggtc tgcaggagtt 42375
 taatctgcat gcgcattctt ttcttgtgac cagtttgatc agtgacggac cgggccgtat 42435
 tcctgcaggc gaaccggtgt actctcatga tgagtgtcaa atgctgggta aacaaggccg 42495
 ggctgatgag gtcattgctgc gctggtcaga ggtccgggag atgcacctca gtggccttgt 42555
 tgagtttcac tcgcacacgc acaccaccg acgctgggac cagaagcctg tgtcccgtaa 42615
 tccgtcggat ttgcttcgtg tcgatattct tcttagtctg aagcggatga gggagatgct 42675
 gggttattgc agtcagcatc tgtgctggcc tgagggctgg tattgttctg actatattca 42735
 tgtggctgaa gagttggggt tcacatacct gtataccaca gaaaggcgta tgaacaatcc 42795
 agtcatcggg tcacagcgta ttggctgtat caacgcaaag gagcgaaaga atgtgggctg 42855
 gctgaaacgt cgtctgtttt atcacaccac gcccggtttt tcttcgctgc tggcccggca 42915
 taagggggca cgtcggatag ctgactgagc cagagaccag gatgaaagtg ctgcataaccg 42975
 aatcctcccc catcatcggc gggcaggggt tgtaggctat atcccaaagt atggtgctga 43035
 ttttgagcag aaataagata acaccggagg ccgtgataga cccggttgcg cagaatctgg 43095
 aggcagcagg tcaactggcg agggcttcta cccgctggct tttggttatg ggagattttg 43155
 aatgtacaga ggcccagcgg gagtggttgt tgttgcccg gaattattgc cttgcgcaga 43215
 tatcttcccc tggcgccagt aaagctggat atcagcgacg tgacgaaatg agaggagcct 43275
 gcacataggt ttgtgttttg agtaagtgtc ttatttaaac cgtctgttct gtttctccg 43335
 ctttcacaaa taatgtcgag ccgggtgggg gactcaagta agaataatct ggcgatgttt 43395
 tgcttgtttc cacgggatac tttgttaggt gaacgatata attaatgcgc tccagactgc 43455

gtcacagaaa aggagaatgc ttttgcagaa cgggcgcagc acaatgaacc actttattat 43515
tggtatctgg ttcggcgaca accttgacga tcacgtctga gatagggtat ttacactgag 43575
tggtaaacag gttcaattag taaccggaga tggatgcaaa atcatgatcg attcagatgc 43635
tattttgcag ccaatagatt ttttattaag atgatatcaa ggacattgag gcacacaatg 43695
acgtatcagt aagtcgttga tagctcattt gatataagaa tttcttttat caacggaaga 43755
taatgatgga actgatcaat aatcgtggta tgcgagactg gatgatattt attaaagtgg 43815
cggaagtagg gaatctttcc cgggctgcgc gggaattaga tattagcatt tctgctgtca 43875
gtaaatcgct tagtcgcctt gagaattcta ttgaggttac tttacttcgg cgcgattcac 43935
atcacttaga actgactgga gctggtcaga cagcctatgc aagcatgaaa aggataacat 43995
cttcttttca gtccttgctg gatgaattgc gaaatccgga taaaattatc agaggaggta 44055
taaaattttc ggctccggct attgtctgtg agtttcttgc caataagtgg atatgggaat 44115
ttacagctag ctatccggat acaaaaatct acctggattc acgagagcgt agcgattttt 44175
ttagtaaatc cctggagttt gatgagctgg tttttaaag tggcataatc gaaagtgagg 44235
atctcgtgta tcgaaagata agccctttaa agttggttct ttgtgcgagt ccgaaatata 44295
tcagaaaata tggcaggatc tcacacctg gcgatttgga aaatcacatt attgtgggtc 44355
ttcacaacca tggcttttcc ggacctotta ctcttttccg tcaggatgaa tcatacacta 44415
ttagtgggcg tgtaaatgtt catttatctt ccaataatct tttgagtgtt cttaatttgg 44475
ttttagaagg aaagggatc aacctcatga ctccggcctg gcttgccacc aaatacttaa 44535
aaaataatga acttgaaatt atacttctg aatggagggt tccagatctc cccatttatt 44595
ttgtatggcg tcatcgtcag tattattctc ctttatttca acgctttctg tcttttattg 44655
aagataaatg gaataatcgc ccacaaattg attttctgaa tgatgattaa cccgtttgga 44715
atggttttga tacgttctg acttaaacc acatgatgac tgaattgagg catcgagata 44775
tgcgactggc cagccagtcg tcttttgacg atgccaata caaacatga tgccgactga 44835
cggaatgat aatacgcgga aacaggacgg ggctgtttt gggcagccgg aagttaagcc 44895
cataccagaa acgttgcagt gtactgaaaa atggcgccag gttgcacctg ttcaaagatt 44955
ttctgaaggc gcaggagtat tcattactga tatctccat tgcgccttcg ggaaccaca 45015
ggaccagcta ttttaccgat agtgtttaaa aggcgtaagt aatgccgagc atgaagtc 45075
tgagggcagc ctttgtgtct gcatcataag cggtatgttc atcaccagca tagtgatttt 45135
ttgaaatgct tactttgcca gcattaatgt atttataact ggcgtcaatc ataattatt 45195
ctgttacagc atattttgca ccgatacctg cgccccaggc aaagttattt tttgaagcag 45255
acagagtttc attaatacca aaaccaacag gaatggtgtt attacttagc ttcacatgag 45315

cgaggccaac gcctgcgctg atatagggag taaatgccgt actattgtga aaatcataat 45375
 agccattaac catgtaagtg gtcattcgga cctgattttt tacatttatg tgtactggat 45435
 caccaaatgc aataatatcc tgcccgccct tagcatccgt ctcacctctg aaagtggat 45495
 ccagttctaa acgtactgga agctggaatg gatcataaaa gtcataaccg atagcaaccc 45555
 cgccgcaaaa aacgcctttg gtacggtcag gtaacgttgc atgaccatta actatctcat 45615
 cctggctgaa ggttgagttg attccataga cattgactac ggatgtcccc gctttcccg 45675
 tgatatagat cccttctttt gctgatgcag tagcggacca ggctaccaca aggggaatga 45735
 tgcagactgc gaaaaagttt ttcatttcag aacctgcctt aatattgggc taaaagacaa 45795
 gtttcacggt ataggggtgtg atataacgat tacataaacg aagcccaaaa aacggtctat 45855
 tgtaacgctg ggttttctgt aagcgggtaa aaaatgagat gaagatttta aataacaata 45915
 cgataatcgt cggtatggaa atccatctcc tcgccaaatt gccccacgta cggtttctact 45975
 tctacgttat gtaacgggta gtgtgagatg gagcgatgct gtaagaaaaa gatgaagatg 46035
 aatttgtagc cgacctggat aaagcccggt atcccggat aacgggcaaa aatatttact 46095
 caagtgcctg ggcgagatct tgtgtacct gttgacgctg ttctgggtgtt aagactttgc 46155
 ttaaataaaa ataataattta acccgataat agcgagcctg ttgttctatg ttactgaagg 46215
 ctgcaagctg ctgttttacg gcggcgctcat cccatttacc ggatttaatc acctctatca 46275
 gcgcaccgtc ttttaattccc ttcatagaaa tctgactgac gtcgggttcc agttgttggt 46335
 gaagtttttt gatccgggta atctgatcgt ttgtcagctt cagatgctgg acaataggat 46395
 cctgggcggg caggggagga ttggggacag cggtggcgaa agcgccaaaa gaaacgccc 46455
 ccagagtcgc tgccagtaaa gttgtgcgta caaagttttt catgaagata tcctgataag 46515
 ggagtgatta accgttttta ttaccacga atggcgagca attatcttag agcctatccc 46575
 agtagggcta ttttacttgc cattttggac ctgggcagtg ctgcgcaaaa cgcgtagcg 46635
 ttttgaacgc cgctagcggc ggcccgaagg gcgagcgtag cgagtcaaac ctcacgtact 46695
 acgtgtacgc tccggttttt gcgcgctgtc cgtgtccaaa ctggctgcgc caataacgcc 46755
 tgggtgggata ggctcttagt cagaatacgt tgccgtccac attacgccac gcgaatttgt 46815
 tttacggaga gttacggagt gaaacaatcc cgccgcggtg agcggcaggt tgctt 46870

<210> 2

<211> 166

<212> PRT

<213> Salmonella typhimurium

<400> 2

Met Lys Ser Ile Lys Lys Leu Ile Ile Ala Ser Ala Leu Ser Met Met

1 5 10 15
 Ala Ala Ser Cys Tyr Ala Gly Ser Phe Leu Pro Asn Ser Glu Gln Gln
 20 25 30
 Lys Ser Val Asp Ile Val Phe Ser Ser Pro Gln Asp Leu Thr Val Ser
 35 40 45
 Leu Ile Pro Val Ser Gly Leu Lys Ala Gly Lys Asn Ala Pro Ser Ala
 50 55 60
 Lys Ile Ala Lys Leu Val Val Asn Ser Thr Thr Leu Lys Glu Phe Gly
 65 70 75 80
 Val Arg Gly Ile Ser Asn Asn Val Val Asp Ser Thr Gly Thr Ala Trp
 85 90 95
 Arg Val Ala Gly Lys Asn Thr Gly Lys Glu Ile Gly Val Gly Leu Ser
 100 105 110
 Ser Asp Ser Leu Arg Arg Ser Asp Ser Thr Glu Lys Trp Asn Gly Val
 115 120 125
 Asn Trp Met Thr Phe Asn Ser Asn Asp Thr Leu Asp Ile Val Leu Thr
 130 135 140
 Gly Pro Ala Gln Asn Val Thr Ala Asp Thr Tyr Pro Ile Thr Leu Asp
 145 150 155 160
 Val Val Gly Tyr Gln Pro
 165

<210> 3
 <211> 245
 <212> PRT
 <213> Salmonella typhimurium

<400> 3
 Met Lys Ile Val Asn Phe Ala Val Met Ala Val Ala Leu Phe Ala Thr
 1 5 10 15
 Asn Ser Met Val Ser Val Tyr Ala Val Asn Gln Gln Leu Asn Ser Ala
 20 25 30
 Thr Lys Leu Phe Ser Val Lys Leu Gly Ala Thr Arg Val Ile Tyr His
 35 40 45
 Ala Gly Thr Ala Gly Ala Thr Leu Ser Val Ser Asn Pro Gln Asn Tyr
 50 55 60
 Pro Ile Leu Val Gln Ser Ser Val Lys Ala Ala Asp Lys Ser Ser Pro
 65 70 75 80
 Ala Pro Phe Leu Val Met Pro Pro Leu Phe Arg Leu Glu Ala Asn Gln
 85 90 95
 Gln Ser Gln Leu Arg Ile Val Arg Thr Gly Gly Asp Met Pro Thr Asp
 100 105 110
 Arg Glu Thr Leu Gln Trp Val Cys Ile Lys Ala Val Pro Pro Glu Asn

115

120

125

Glu Pro Ser Asp Thr Gln Ala Lys Gly Ala Thr Leu Asp Leu Asn Leu
130 135 140

Ser Ile Asn Ala Cys Asp Lys Leu Ile Phe Arg Pro Asp Ala Val Lys
145 150 155 160

Gly Thr Pro Glu Asp Val Ala Gly Asn Leu Arg Trp Val Glu Thr Gly
165 170 175

Asn Lys Leu Lys Val Glu Asn Pro Thr Pro Phe Tyr Met Asn Leu Ala
180 185 190

Ser Val Thr Val Gly Gly Lys Pro Ile Thr Gly Leu Glu Tyr Val Pro
195 200 205

Pro Phe Ala Asp Lys Thr Leu Asn Met Pro Gly Ser Ala His Gly Asp
210 215 220

Ile Glu Trp Arg Val Ile Thr Asp Phe Gly Gly Glu Ser His Pro Phe
225 230 235 240

His Tyr Val Leu Lys
245

<210> 4

<211> 836

<212> PRT

<213> Salmonella typhimurium

<400> 4

Met Lys Phe Lys Gln Pro Ala Leu Leu Leu Phe Ile Ala Gly Val Val
1 5 10 15

His Cys Ala Asn Ala His Thr Tyr Thr Phe Asp Ala Ser Met Leu Gly
20 25 30

Asp Ala Ala Lys Gly Val Asp Met Ser Leu Phe Asn Gln Gly Leu Gln
35 40 45

Gln Pro Gly Thr Tyr Arg Val Asp Val Met Val Asn Gly Lys Arg Val
50 55 60

Asp Thr Arg Asp Val Val Phe Lys Leu Glu Lys Asp Gly Gln Gly Thr
65 70 75 80

Pro Val Leu Ala Pro Cys Leu Thr Val Ser Gln Leu Ser Arg Tyr Gly
85 90 95

Val Lys Thr Glu Asp Tyr Pro Gln Leu Trp Lys Ala Ala Lys Pro Pro
100 105 110

Asp Glu Cys Ala Asp Leu Thr Ala Ile Pro Gln Ala Lys Ala Val Leu
115 120 125

Asp Ile Asn Asn Gln Gln Leu Gln Leu Ser Ile Pro Gln Leu Ala Leu
130 135 140

Arg Pro Glu Phe Lys Gly Ile Ala Pro Glu Asp Leu Trp Asp Asp Gly

145													155													160
Ile	Pro	Ala	Phe	Leu	Met	Asn	Tyr	Ser	Ala	Arg	Thr	Thr	Gln	Thr	Asp											
				165					170						175											
Tyr	Lys	Met	Asp	Met	Val	Gly	Arg	Asp	Asn	Ser	Ser	Trp	Val	Gln	Leu											
			180					185					190													
Gln	Pro	Gly	Ile	Asn	Ile	Gly	Ala	Trp	Arg	Val	Arg	Asn	Ala	Thr	Ser											
		195					200					205														
Trp	Gln	Arg	Ser	Ser	Gln	Leu	Ser	Gly	Lys	Trp	Gln	Ala	Ala	Tyr	Thr											
	210					215					220															
Tyr	Ala	Glu	Arg	Gly	Leu	Tyr	Ser	Leu	Lys	Ser	Arg	Leu	Thr	Leu	Gly											
225					230					235					240											
Gln	Lys	Thr	Ser	Gln	Gly	Glu	Ile	Phe	Asp	Ser	Val	Pro	Phe	Thr	Gly											
				245					250					255												
Val	Met	Leu	Ala	Ser	Asp	Asp	Asn	Met	Val	Pro	Tyr	Ser	Glu	Arg	Gln											
			260					265					270													
Phe	Ala	Pro	Val	Val	Arg	Gly	Ile	Ala	Arg	Thr	Gln	Ala	Arg	Val	Glu											
		275					280					285														
Val	Lys	Gln	Asn	Gly	Tyr	Thr	Ile	Tyr	Asn	Thr	Thr	Val	Ala	Pro	Gly											
	290					295					300															
Pro	Phe	Ala	Leu	Arg	Asp	Leu	Ser	Val	Thr	Asp	Ser	Ser	Gly	Asp	Leu											
305					310					315					320											
His	Val	Thr	Val	Trp	Glu	Ala	Asp	Gly	Ser	Thr	Gln	Met	Phe	Val	Val											
				325					330					335												
Pro	Tyr	Gln	Thr	Pro	Ala	Ile	Ala	Leu	His	Gln	Gly	Tyr	Leu	Lys	Tyr											
			340					345					350													
Ser	Leu	Leu	Ala	Gly	Arg	Tyr	Arg	Ser	Ser	Asp	Ser	Ala	Thr	Asp	Lys											
		355					360					365														
Arg	Gln	Ile	Ala	Gln	Ala	Thr	Leu	Met	Tyr	Gly	Leu	Pro	Trp	Asn	Leu											
		370				375					380															
Thr	Ala	Tyr	Gly	Gly	Ile	Gln	Ser	Ala	Thr	His	Asn	Gln	Ala	Ala	Leu											
385					390					395					400											
Leu	Gly	Leu	Gly	Gly	Ser	Leu	Gly	Arg	Trp	Gly	Ser	Leu	Ser	Val	Asp											
				405					410					415												
Gly	Ser	Asp	Thr	His	Ser	Gln	Arg	Gln	Gly	Glu	Ala	Val	Gln	Gln	Gly											
			420					425					430													
Ala	Ser	Trp	Arg	Leu	Arg	Tyr	Ser	Asn	Gln	Leu	Thr	Ala	Thr	Gly	Thr											
		435					440					445														
Asn	Phe	Phe	Leu	Thr	Arg	Trp	Gln	Tyr	Ala	Ser	Gln	Gly	Tyr	Asn	Thr											
	450					455					460															
Leu	Ser	Asp	Val	Leu	Asp	Ser	Tyr	Arg	His	Asn	Gly	Asn	Arg	Leu	Trp											
465					470					475					480											

Ser Trp Arg Glu Asn Leu Gln Pro Ser Ser Arg Thr Thr Leu Met Leu
 485 490 495
 Ser Gln Ser Trp Gly Arg His Leu Gly Asn Leu Ser Leu Thr Gly Ser
 500 505 510
 Arg Thr Asp Trp Arg Asn Arg Pro Gly His Asp Asp Ser Tyr Gly Leu
 515 520 525
 Ser Trp Gly Thr Ser Ile Gly Gly Gly Ser Leu Ser Leu Asn Trp Asn
 530 535 540
 Gln Asn Arg Thr Leu Trp Arg Asn Gly Ala His Arg Lys Glu Asn Ile
 545 550 555 560
 Thr Ser Leu Trp Phe Ser Met Pro Leu Ser Arg Trp Thr Gly Asn Asn
 565 570 575
 Val Ser Ala Ser Trp Gln Met Thr Ser Pro Ser His Gly Gly Gln Thr
 580 585 590
 Gln Gln Val Gly Val Asn Gly Glu Ala Phe Ser Gln Gln Leu Asp Trp
 595 600 605
 Glu Val Arg Gln Ser Tyr Arg Ala Asp Ala Pro Pro Gly Gly Gly Asn
 610 615 620
 Asn Ser Ala Leu His Leu Ala Trp Asn Gly Asp Tyr Gly Leu Leu Gly
 625 630 635 640
 Gly Asp Tyr Ser Tyr Ser Arg Ala Met Arg Gln Met Gly Val Asn Ile
 645 650 655
 Ala Gly Gly Ile Val Ile His His His Gly Val Thr Leu Gly Gln Pro
 660 665 670
 Leu Gln Gly Ser Val Ala Leu Val Glu Ala Pro Gly Ala Ser Gly Val
 675 680 685
 Pro Val Gly Gly Trp Pro Gly Val Lys Thr Asp Phe Arg Gly Asp Thr
 690 695 700
 Thr Val Gly Asn Leu Asn Val Tyr Gln Glu Asn Thr Val Ser Leu Asp
 705 710 715 720
 Pro Ser Arg Leu Pro Asp Asp Ala Glu Val Thr Gln Thr Asp Val Arg
 725 730 735
 Val Val Pro Thr Glu Gly Ala Val Val Glu Ala Lys Phe His Thr Arg
 740 745 750
 Ile Gly Ala Arg Ala Leu Met Thr Leu Lys Arg Glu Asp Gly Ser Ala
 755 760 765
 Ile Pro Phe Gly Ala Gln Val Thr Val Asn Gly Gln Asp Gly Ser Ala
 770 775 780
 Ala Leu Val Asp Thr Asp Ser Gln Val Tyr Leu Thr Gly Leu Ala Asp
 785 790 795 800

Lys Gly Glu Leu Thr Val Lys Trp Gly Ala Gln Gln Cys Arg Val Asn
 805 810 815

Tyr Arg Leu Pro Ala His Lys Gly Ile Ala Gly Leu Tyr Gln Met Ser
 820 825 830

Gly Leu Cys Arg
 835

<210> 5

<211> 156

<212> PRT

<213> Salmonella typhimurium

<400> 5

Met Trp Met Lys Ile Gln Arg Val Lys Thr Val Ile Tyr Ser Val Ser
 1 5 10 15

Leu Leu Val Ala Ala Ser Ser Leu Val Pro Ile Ala Asn Ala Ala Glu
 20 25 30

Lys Leu Gln Thr Thr Leu Arg Val Gly Thr Tyr Phe Arg Ala Gly His
 35 40 45

Val Pro Asp Gly Met Val Leu Ala Gln Gly Trp Val Thr Tyr His Gly
 50 55 60

Ser His Ser Gly Phe Arg Val Trp Ser Asp Glu Gln Lys Ala Gly Asn
 65 70 75 80

Thr Pro Thr Val Leu Leu Leu Ser Gly Gln Gln Asp Pro Arg His His
 85 90 95

Ile Gln Val Arg Leu Glu Gly Glu Gly Trp Gln Pro Asp Thr Val Ser
 100 105 110

Gly Arg Gly Ala Ile Leu Arg Thr Ala Ala Asp Asn Ala Ser Phe Ser
 115 120 125

Val Val Val Asp Gly Asn Gln Glu Val Pro Ala Asp Thr Trp Thr Leu
 130 135 140

Asp Phe Lys Ala Cys Ala Leu Ala Gln Glu Asp Thr
 145 150 155

SEQUENCE LISTING NO. 2

<110> Folkesson, Anders

<120> The complete sequence of the tcf insert of Salmonella enterica serovar Typhi.

<130> The tcf insert in Salmonella typhi

<140>

<141>

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 9253

<212> DNA

<213> Salmonella typhi

<220>

<221> CDS

<222> (1898)..(2608)

<223> tcfA putative fimbrial subunit

<220>

<221> CDS

<222> (2659)..(3234)

<223> tcfB putative fimbrial subunit

<220>

<221> CDS

<222> (3360)..(6029)

<223> tcfC putative fimbrial subunit

<220>

<221> CDS

<222> (6052)..(7131)

<223> tcfD putative fimbrial subunit

<220>

<221> CDS

<222> (7264)..(7719)

<223> tinR putative transcriptional regulator

<400> 1

```

tgtaagtatc ccgcataatc gagccattca catttagaga tcatccggca taatcaatct 60
gccaacgcag gagatcgctg tgcgtaaagc ccgtattact gcgcaccaga tcatcgctgt 120
gattagatca gttgaatccg gacggactgt taaagatgtc taccgggagg ccggtatttc 180
tgaagccacc agggacaact ggaagtcttg atacggcggc atggaagctt ctgatattaa 240
atcttgagga tgtcaacgcc aggatttatg gtcgtttttc gctattttta atatccgctg 300
tttgatcact tctgctgtcc gctttccgcc atttcatctt cactgattgg cgttgcgctt 360
ttggtcagcg ccccgacttt gcgtttttcc ttcagttggt aactttctcc tttgatattc 420

```

agtgtggttg agt gatgcag tagccgatcc aggatcgctg ttgccagcac gttatcaccg 480
 aacatctctc cccagtctgc gaagcctttg tttgacgtca ggataatgct cgttttttca 540
 taccgtcggg tcagcagtcg gaagaacaga ctggcttcct cactgggcat tggcaggtaa 600
 cctatttcgt ccaggatcag cccccgcgca tagctcagtt gttgtagctg gcgctccagg 660
 cggttttcca gtttcgcttt catcagcgtc taccaacctg tcctgcggca tgaacaacac 720
 ccgatggccc acatctgccg ctttcacacc gggggcagcg gccagggtggg ttttctccac 780
 tccggggggc ccagcaggat cacattctcg cagcgtcca cgaacgccag accggccaac 840
 tcccggacga ccttacgatc gatgcctggc tggaagctga agtcgaactg ctccagcgtt 900
 ttgactcatg gcagacgagc ctgtttcagc cgggattcca ttccgcgctg atgtctgccg 960
 ttctattcct gctgcagcg catgcacagg aatgtcgact tat ttttgggc tgacagaggc 1020
 attggctcgtt aagacttgta tattcatgtt gcttactcgc aaaaagggg gaaaggaacc 1080
 gtaacaaca ttgacaggta agggactggc atcaatttga atggtat ttttataacatt 1140
 atgttttaaa aggtgaattg tttttcatg gtgagttgtt ctgttttata ttgttggttg 1200
 tggaatattt agctaattt taatctaaaa atagtgtatt atattgatta cactttgctg 1260
 gagagggtga aacatcaagc gtaaccatat gatgtatata agtttttggt tgctgatatc 1320
 atatttttaa cagatat tga gcaagtaa ctgatctaac tgtaacttta gttttctaaa 1380
 ttaattactt tgcttacttt tttttacctg gttttgtaa acctcatgat gatcagtcta 1440
 tttgtggctt tgtttaaggt ttatttggtt ttgtagtctg aattgtatct cctcctgga 1500
 gggctttctt tagattctct gtctcctacg ttgttatgtt acgtatttgt tgctttgaag 1560
 gagggggaaa tacagttcca tttatctgag taagtcagg acacagtaac aactttctta 1620
 tgaagaattt ccaaaatttt tactgcggcg ttattaattg ttcagcgatt cttacagatc 1680
 tgcgttcgc ttttggtgaa tgaaatccgt ggacttttat ttactaattt tttctttcct 1740
 gaaaaaaca gaggtattga gcgaaaaatt ttattccgta tgatgccctc cacacaaaat 1800
 gtattaacac tgaatcgtaa tttgcttctt tatgctgata actttctgtc tatgctaata 1860
 ctaaaattta gatgactttt atacggtaaa atctgggt atg aat ttt aaa gat act 1915
 Met Asn Phe Lys Asp Thr
 1 5
 ctt ccc ggg gtg ttt ctc tgt gtc gct atg ttt gca tgt ggt cat gcc 1963
 Leu Pro Gly Val Phe Leu Cys Val Ala Met Phe Ala Cys Gly His Ala
 10 15 20
 agg gcg aat atg ctc gtt tat ccc atg gcg gca gaa att aat agt agc 2011
 Arg Ala Asn Met Leu Val Tyr Pro Met Ala Ala Glu Ile Asn Ser Ser
 25 30 35

cgc gaa gag gcc acc tcg ctg ttc gtc tat tct aaa tca gat cat gtg 2059
 Arg Glu Glu Ala Thr Ser Leu Phe Val Tyr Ser Lys Ser Asp His Val
 40 45 50

caa tat att cga aca aga atc atg cgt att gaa cac ccc ggt atg cca 2107
 Gln Tyr Ile Arg Thr Arg Ile Met Arg Ile Glu His Pro Gly Met Pro
 55 60 65 70

cag gag aag gag gta cca gca ggg aat gat ata gag aca gga ctt gtt 2155
 Gln Glu Lys Glu Val Pro Ala Gly Asn Asp Ile Glu Thr Gly Leu Val
 75 80 85

gtc tcc ccg gag aaa ttt gct ctt tcc ccg gga aca aaa aaa aca ata 2203
 Val Ser Pro Glu Lys Phe Ala Leu Ser Pro Gly Thr Lys Lys Thr Ile
 90 95 100

cgt gtt atc agt act cag gca ccg gaa aga gag gaa gcc tgg cgg gta 2251
 Arg Val Ile Ser Thr Gln Ala Pro Glu Arg Glu Glu Ala Trp Arg Val
 105 110 115

tac ttc gag gct gtt cct gaa ctg gaa gat gat cca cag gca ggc gga 2299
 Tyr Phe Glu Ala Val Pro Glu Leu Glu Asp Asp Pro Gln Ala Gly Gly
 120 125 130

aag caa aat tca tcc gta agt gtg aat ctt gtc tgg ggg gtg ttg ctg 2347
 Lys Gln Asn Ser Ser Val Ser Val Asn Leu Val Trp Gly Val Leu Leu
 135 140 145 150

cgt gtt tct ccg tca gac ccc agg cct gcg ctg gta acg gac ggt cac 2395
 Arg Val Ser Pro Ser Asp Pro Arg Pro Ala Leu Val Thr Asp Gly His
 155 160 165

cac ctg ctg aat acg gga aac aca cgg ctt tct ctt att cgg gct ggc 2443
 His Leu Leu Asn Thr Gly Asn Thr Arg Leu Ser Leu Ile Arg Ala Gly
 170 175 180

aac tgc gac acc aca tgc cac tgg cag aat ata ggc aaa agt att tat 2491
 Asn Cys Asp Thr Thr Cys His Trp Gln Asn Ile Gly Lys Ser Ile Tyr
 185 190 195

ccc ggc ggg agt gct gat att ccg gcc gga ata aaa agt aat gca ttt 2539
 Pro Gly Gly Ser Ala Asp Ile Pro Ala Gly Ile Lys Ser Asn Ala Phe
 200 205 210

cgt gtg gaa tat cgt acg ggt gca aat tca ccg gta atc tct gct gat 2587
 Arg Val Glu Tyr Arg Thr Gly Ala Asn Ser Pro Val Ile Ser Ala Asp
 215 220 225 230

tta aca gca gcc gga aag taa aaacacacgg agcgtacgct atacctaca 2638
 Leu Thr Ala Ala Gly Lys
 235

tttattctca gggggagcgg atg tat acc gag tgt aca tat atc act gta ata 2691
 Met Tyr Thr Glu Cys Thr Tyr Ile Thr Val Ile
 240 245

aac aac aaa gca agg tta ttt ttt atg aac atg aaa aca tct ttt att 2739
 Asn Asn Lys Ala Arg Leu Phe Phe Met Asn Met Lys Thr Ser Phe Ile

250	255	260	
gcc gca gct gtg gca ttg gcc acc gtt tat tct ttt tct gtt tct gcg			2787
Ala Ala Ala Val Ala Leu Ala Thr Val Tyr Ser Phe Ser Val Ser Ala			
265	270	275	280
gtt cag aag gat att acc gtc act gcc aat att gac agt aca ctt gaa			2835
Val Gln Lys Asp Ile Thr Val Thr Ala Asn Ile Asp Ser Thr Leu Glu			
	285	290	295
ctg ctg cag gcc gat ggt tca tcc ctc ccg tcg act atg aag ctg gat			2883
Leu Leu Gln Ala Asp Gly Ser Ser Leu Pro Ser Thr Met Lys Leu Asp			
	300	305	310
ttc atg ccg ggt aag ggc ctg gtc cat aaa tca ctc cag acc cgc ctt			2931
Phe Met Pro Gly Lys Gly Leu Val His Lys Ser Leu Gln Thr Arg Leu			
	315	320	325
tac agc aac gat cag acc aag tcg gtt aat gta aaa ctg ttg aat gct			2979
Tyr Ser Asn Asp Gln Thr Lys Ser Val Asn Val Lys Leu Leu Asn Ala			
	330	335	340
cca caa ctt atc aac gtc ctg gat ccc acc aaa acc att gat atg gaa			3027
Pro Gln Leu Ile Asn Val Leu Asp Pro Thr Lys Thr Ile Asp Met Glu			
	345	350	355
gtg act ctg gga gga cgg tca ctg acc acc acc aat tct gta ctg gaa			3075
Val Thr Leu Gly Gly Arg Ser Leu Thr Thr Thr Asn Ser Val Leu Glu			
	365	370	375
gct aaa acc ctg ttc ccg gac gga aaa act ggc gat gct tca gct ctg			3123
Ala Lys Thr Leu Phe Pro Asp Gly Lys Thr Gly Asp Ala Ser Ala Leu			
	380	385	390
ctg aac ctg gat att ggt cag aag gct gga gca gcc tta caa aac ctg			3171
Leu Asn Leu Asp Ile Gly Gln Lys Ala Gly Ala Ala Leu Gln Asn Leu			
	395	400	405
cct gcc ggt gaa tac agc gga ttg gtc agt ctg gtg att tca cag gct			3219
Pro Ala Gly Glu Tyr Ser Gly Leu Val Ser Leu Val Ile Ser Gln Ala			
	410	415	420
gtc act gcc ggc taa taactgggta ttagctcttc atctgatccg gttttggggg			3274
Val Thr Ala Gly			
425			
gcaccgttcg tacctgaacc ggatccggta ttgatcttat tattcattgc aattcaggtc			3334
tctttacgtg agtcgttatt tctgg atg tat tat tta ctg gga ttg tgc agt			3386
		Met Tyr Tyr Leu Leu Gly Leu Cys Ser	
	430	435	
ttt acc agc cag gca act ctt att ccc cct cct gga ttt gaa tct ctg			3434
Phe Thr Ser Gln Ala Thr Leu Ile Pro Pro Pro Gly Phe Glu Ser Leu			
	440	445	450
ctg gaa gga cag act gag caa att gaa gtg ttg cta cca ggg cat tca			3482
Leu Glu Gly Gln Thr Glu Gln Ile Glu Val Leu Leu Pro Gly His Ser			
	455	460	465
			470

ctg gga tta ttt ccg gtg gtg gtt aaa ccg gac acc gtg cag ttc atg Leu Gly Leu Phe Pro Val Val Val Lys Pro Asp Thr Val Gln Phe Met 475 480 485	3530
tcc cca ttg atg gta ctt gaa agc agt ggg ctt gcc gcg ttg ccg gcc Ser Pro Leu Met Val Leu Glu Ser Ser Gly Leu Ala Ala Leu Pro Ala 490 495 500	3578
gca gaa cgg caa aaa gcg ctg gct gca ctc agc cgt ccg ttg cta cgt Ala Glu Arg Gln Lys Ala Leu Ala Ala Leu Ser Arg Pro Leu Leu Arg 505 510 515	3626
aac agc aat ctg gtc tgt ggt gtc tca gaa gca aaa gac agc agc gag Asn Ser Asn Leu Val Cys Gly Val Ser Glu Ala Lys Asp Ser Ser Glu 520 525 530	3674
tgt ggt tac gtg gca aca gat aaa gag gat gtt gcg gtt att ttt gat Cys Gly Tyr Val Ala Thr Asp Lys Glu Asp Val Ala Val Ile Phe Asp 535 540 545 550	3722
gag aac aac gct cag tta tct ttg ttt ctt aac cgg gac tgg ttg ccg Glu Asn Asn Ala Gln Leu Ser Leu Phe Leu Asn Arg Asp Trp Leu Pro 555 560 565	3770
gat gaa gaa cga cgt gat aaa cgc tgg ctg act ccg acc ccg gag ggt Asp Glu Glu Arg Arg Asp Lys Arg Trp Leu Thr Pro Thr Pro Glu Gly 570 575 580	3818
gtc agc gca ttt att cac cgc cag acg ctg tat ctg agt gat gat ctc Val Ser Ala Phe Ile His Arg Gln Thr Leu Tyr Leu Ser Asp Asp Leu 585 590 595	3866
cac agt cgt aat atg aca ctg aat ggt agc ggt gcc ctg ggg ctt ggt His Ser Arg Asn Met Thr Leu Asn Gly Ser Gly Ala Leu Gly Leu Gly 600 605 610	3914
gac ggt cgt tat ctg gga ggc gac tgg gcg gct atc tgg aat cag tca Asp Gly Arg Tyr Leu Gly Gly Asp Trp Ala Ala Ile Trp Asn Gln Ser 615 620 625 630	3962
gaa cat tac aat aac agt cag gcc tgg ttt gac aat ctg ttt gtc cgt Glu His Tyr Asn Asn Ser Gln Ala Trp Phe Asp Asn Leu Phe Val Arg 635 640 645	4010
cag gat ctc ggc aat cag tat tat ctc cag gct ggt cgg atg gat cag Gln Asp Leu Gly Asn Gln Tyr Tyr Leu Gln Ala Gly Arg Met Asp Gln 650 655 660	4058
cgg aat ctg tcc agc gcc acg ggg ggg gat ttt ggg ttc agt ctg ctt Arg Asn Leu Ser Ser Ala Thr Gly Gly Asp Phe Gly Phe Ser Leu Leu 665 670 675	4106
ccc ctg agc cgg ttt gat gga tta cga acc ggg acc acc caa gct tat Pro Leu Ser Arg Phe Asp Gly Leu Arg Thr Gly Thr Thr Gln Ala Tyr 680 685 690	4154
gtt aac cat gag gtg gac cat aat gcc act ccg gtt atg gtt cag gtt Val Asn His Glu Val Asp His Asn Ala Thr Pro Val Met Val Gln Val 695 700 705 710	4202

acc cga aat gcc cgt att gat att tat cgt ggc agc gag ttg ctg ggg	4250
Thr Arg Asn Ala Arg Ile Asp Ile Tyr Arg Gly Ser Glu Leu Leu Gly	
715 720 725	
agt cag ttc ctg acc ccg gga atg cat acc ctg gat act cat tct ctt	4298
Ser Gln Phe Leu Thr Pro Gly Met His Thr Leu Asp Thr His Ser Leu	
730 735 740	
cca ccg gga agc tat cct ctg gcg ttg cgg gtg tat gag gat ggg att	4346
Pro Pro Gly Ser Tyr Pro Leu Ala Leu Arg Val Tyr Glu Asp Gly Ile	
745 750 755	
ctg cgg cga acg gag acc cag ccc ttc agt aag ggg ggc aat agc ttc	4394
Leu Arg Arg Thr Glu Thr Gln Pro Phe Ser Lys Gly Gly Asn Ser Phe	
760 765 770	
agt gca cag acc cag tgg ttt att cag ggc ggg ctg gaa gat acc ggg	4442
Ser Ala Gln Thr Gln Trp Phe Ile Gln Gly Gly Leu Glu Asp Thr Gly	
775 780 785 790	
gat aaa gcc agc cat tat gac ggt gag act gtc atg gct gcc gga ttc	4490
Asp Lys Ala Ser His Tyr Asp Gly Glu Thr Val Met Ala Ala Gly Phe	
795 800 805	
caa act ggg ctg cgg aaa aat atc agt ctg acc gaa ggt atc tct ctg	4538
Gln Thr Gly Leu Arg Lys Asn Ile Ser Leu Thr Glu Gly Ile Ser Leu	
810 815 820	
gca cat gag gcc tgg tac agt gaa acc cga ctg aat tca cag cat gca	4586
Ala His Glu Ala Trp Tyr Ser Glu Thr Arg Leu Asn Ser Gln His Ala	
825 830 835	
gtg ctg gat ggc acg ctg gac ctt tct gcc ggg ata ctg cat ggg aca	4634
Val Leu Asp Gly Thr Leu Asp Leu Ser Ala Gly Ile Leu His Gly Thr	
840 845 850	
gac agc acg agc ggt aac act gag cag gtg aca tac aac gac gga ttt	4682
Asp Ser Thr Ser Gly Asn Thr Glu Gln Val Thr Tyr Asn Asp Gly Phe	
855 860 865 870	
tcc gcg agt ctg tgg cgt aac cat acg gaa agt gat gcc tgt agt ggt	4730
Ser Ala Ser Leu Trp Arg Asn His Thr Glu Ser Asp Ala Cys Ser Gly	
875 880 885	
cgt cat cca cag tca gtg cat gcc agt atg acc tgc cag act tcg atg	4778
Arg His Pro Gln Ser Val His Ala Ser Met Thr Cys Gln Thr Ser Met	
890 895 900	
aac gcc tcc ctg tcg gtt tcg gtg ggg aac tgg tat gcc cta ctg gga	4826
Asn Ala Ser Leu Ser Val Ser Val Gly Asn Trp Tyr Ala Leu Leu Gly	
905 910 915	
tac agt acc agc agg aca gaa ggt cgg ccg gtt tac cgg gga tat gat	4874
Tyr Ser Thr Ser Arg Thr Glu Gly Arg Pro Val Tyr Arg Gly Tyr Asp	
920 925 930	
gat aac agt gac aaa gaa aat gtg ttc tgg cga cag gca tac atc cct	4922
Asp Asn Ser Asp Lys Glu Asn Val Phe Trp Arg Gln Ala Tyr Ile Pro	
935 940 945 950	

gcc tct cac cgc gaa tct gct cag gct agt gca acg tac agc ctt aat 4970
 Ala Ser His Arg Glu Ser Ala Gln Ala Ser Ala Thr Tyr Ser Leu Asn
 955 960 965

atg gct ggc atg aat att aat acc cat ggg gga gta tgg cga acc cga 5018
 Met Ala Gly Met Asn Ile Asn Thr His Gly Gly Val Trp Arg Thr Arg
 970 975 980

aat gac gga gtg aat gat gat ggc ttg ttt atg agt gtc agt gtg tca 5066
 Asn Asp Gly Val Asn Asp Asp Gly Leu Phe Met Ser Val Ser Val Ser
 985 990 995

tat gcc tct caa cca ccg aca atg act ggc agt aat agg tat acc tca 5114
 Tyr Ala Ser Gln Pro Pro Thr Met Thr Gly Ser Asn Arg Tyr Thr Ser
 1000 1005 1010

gcc ggg acc gat att cac agt agc cgg aat caa aaa aca cag acg tcc 5162
 Ala Gly Thr Asp Ile His Ser Ser Arg Asn Gln Lys Thr Gln Thr Ser
 1015 1020 1025 1030

tgg aat gtg aac cat gtg aga tcc tgg cag cag gat ctg tat cgt gaa 5210
 Trp Asn Val Asn His Val Arg Ser Trp Gln Gln Asp Leu Tyr Arg Glu
 1035 1040 1045

ctg tcg gtg ggt ttc tcc ggt tat aac gac gac agc tgg agc ggg agt 5258
 Leu Ser Val Gly Phe Ser Gly Tyr Asn Asp Asp Ser Trp Ser Gly Ser
 1050 1055 1060

ctc ggc gga cgc atg agc ggc cgt atg ggt gaa ctg agc gcc act atc 5306
 Leu Gly Gly Arg Met Ser Gly Arg Met Gly Glu Leu Ser Ala Thr Ile
 1065 1070 1075

agt aac tcc cat caa cgt aat gcg ggc agc gcc agt tca ctc acc gct 5354
 Ser Asn Ser His Gln Arg Asn Ala Gly Ser Ala Ser Ser Leu Thr Ala
 1080 1085 1090

ggc tac agc tcg tct ctg gcg tta tcc cgt aat gga ctg ttc tgg gga 5402
 Gly Tyr Ser Ser Ser Leu Ala Leu Ser Arg Asn Gly Leu Phe Trp Gly
 1095 1100 1105 1110

ggt ggt cag gac ggt gaa ccg gcc tct ggc atg gcg gtg aac gtg gag 5450
 Gly Gly Gln Asp Gly Glu Pro Ala Ser Gly Met Ala Val Asn Val Glu
 1115 1120 1125

tca gag ggg gac gag ggc agt agc ggg aaa gta gtc agc gtt cgt ggc 5498
 Ser Glu Gly Asp Glu Gly Ser Ser Gly Lys Val Val Ser Val Arg Gly
 1130 1135 1140

agc agc cag ccg ttc agt ctc ggt ttt ggt cag cag tcg ctg ttg ctg 5546
 Ser Ser Gln Pro Phe Ser Leu Gly Phe Gly Gln Gln Ser Leu Leu Leu
 1145 1150 1155

atg gaa ggc tat aac gcc acg gag gtg acc att gag gat gca ggg gtt 5594
 Met Glu Gly Tyr Asn Ala Thr Glu Val Thr Ile Glu Asp Ala Gly Val
 1160 1165 1170

agt tca cag ggt atg gca ggc gta aaa gcg gga ggg gga agc agg tgt 5642
 Ser Ser Gln Gly Met Ala Gly Val Lys Ala Gly Gly Gly Ser Arg Cys

1175	1180	1185	1190	
tac ttc ctg aca ccc ggg cat ctg ctg gtt cac aac atc agc gcc agt				5690
Tyr Phe Leu Thr Pro Gly His Leu Leu Val His Asn Ile Ser Ala Ser				
1195	1200	1205		
atg agc cga ctg tac gtt ggc cgc gta ctg gac aag gat ggc aga ccg				5738
Met Ser Arg Leu Tyr Val Gly Arg Val Leu Asp Lys Asp Gly Arg Pro				
1210	1215	1220		
ctg ctg gac gca cag cca ctg aac tat cca ttt ttg tgc ttg gga cct				5786
Leu Leu Asp Ala Gln Pro Leu Asn Tyr Pro Phe Leu Ser Leu Gly Pro				
1225	1230	1235		
tcc ggg cga ttt agc ctg cag agc gag cat aaa gaa tcc agc ctg tgg				5834
Ser Gly Arg Phe Ser Leu Gln Ser Glu His Lys Glu Ser Ser Leu Trp				
1240	1245	1250		
ctg ctg tct aaa aac agg atc ctg cgt tgt ccg atg tca gta cat aaa				5882
Leu Leu Ser Lys Asn Arg Ile Leu Arg Cys Pro Met Ser Val His Lys				
1255	1260	1265	1270	
cgt cgg gat gtt atg cag gta gtg ggt gat gtg cgg tgt gaa tta agt				5930
Arg Arg Asp Val Met Gln Val Val Gly Asp Val Arg Cys Glu Leu Ser				
1275	1280	1285		
gac gtg gat gcc ctg cca cag gcg ttg caa ata tgc ccg cgg gtc atc				5978
Asp Val Asp Ala Leu Pro Gln Ala Leu Gln Ile Ser Pro Arg Val Ile				
1290	1295	1300		
cgt ttg ctg aac gtg gca ggt ttg ctg cgc cat tcc gtt cag gaa gcc				6026
Arg Leu Leu Asn Val Ala Gly Leu Leu Arg His Ser Val Gln Glu Ala				
1305	1310	1315		
tga cgtagagata aaggcggttaa ct atg agt aat aaa atg aag tgg acg agt				6078
Met Ser Asn Lys Met Lys Trp Thr Ser				
1320	1325			
atg aca gcc cat tgg tca gca att att aat ttc atc cga aaa tat gtt				6126
Met Thr Ala His Trp Ser Ala Ile Ile Asn Phe Ile Arg Lys Tyr Val				
1330	1335	1340		
tat cca gca agg ata att gcc atc ctg ctg atg gct ggc gct aca ctg				6174
Tyr Pro Ala Arg Ile Ile Ala Ile Leu Leu Met Ala Gly Ala Thr Leu				
1345	1350	1355	1360	
cca caa gtc gcc gat gcg att acc gtc gac ctg aat tac gac aag aac				6222
Pro Gln Val Ala Asp Ala Ile Thr Val Asp Leu Asn Tyr Asp Lys Asn				
1365	1370	1375		
aat gta gcg gtc atc act cct gtc tgg tcc caa gaa tgg agt gta gca				6270
Asn Val Ala Val Ile Thr Pro Val Trp Ser Gln Glu Trp Ser Val Ala				
1380	1385	1390		
aat gtg ttg ggg gga tgg gta tgt cgt tca aac agg aat gaa aat gag				6318
Asn Val Leu Gly Gly Trp Val Cys Arg Ser Asn Arg Asn Glu Asn Glu				
1395	1400	1405		
ggg gcg tgt gaa gaa aca cat ttg gta tgg tgg tat gct ttt gga gct				6366
Gly Ala Cys Glu Glu Thr His Leu Val Trp Trp Tyr Ala Phe Gly Ala				

1410	1415	1420	
tat tca aaa att cgt ctg cgt ttc aga gaa caa atc agc cat gcc gaa			6414
Tyr Ser Lys Ile Arg Leu Arg Phe Arg Glu Gln Ile Ser His Ala Glu			
1425	1430	1435	1440
att acg ctc ata ctg ctc ggc agt gtt cgt gat gcc tgt tat act ggt			6462
Ile Thr Leu Ile Leu Leu Gly Ser Val Arg Asp Ala Cys Tyr Thr Gly			
1445	1450		1455
gtc atc aac atg aac gct gct gca tgt caa tgg ggt agg tcg ctg aaa			6510
Val Ile Asn Met Asn Ala Ala Ala Cys Gln Trp Gly Arg Ser Leu Lys			
1460	1465		1470
ctt agg ata cct tca gaa gag ctt gcg aag ata cct aca agc gga aca			6558
Leu Arg Ile Pro Ser Glu Glu Leu Ala Lys Ile Pro Thr Ser Gly Thr			
1475	1480		1485
tgg aaa gca acg tta gtc ctg gat tat tta caa tgg ggc gga gac gat			6606
Trp Lys Ala Thr Leu Val Leu Asp Tyr Leu Gln Trp Gly Gly Asp Asp			
1490	1495		1500
cct tta ggc aca tca act aca gat atc acg ctg aat gta aca gac cac			6654
Pro Leu Gly Thr Ser Thr Thr Asp Ile Thr Leu Asn Val Thr Asp His			
1505	1510	1515	1520
ttt gct gaa aat gcg gct att tac ttt ccg caa ttt ggt aca gca acg			6702
Phe Ala Glu Asn Ala Ala Ile Tyr Phe Pro Gln Phe Gly Thr Ala Thr			
1525	1530		1535
ccc cgg gtg gac ctg aat ctt cac cgg atg aat gcc tca caa atg tcg			6750
Pro Arg Val Asp Leu Asn Leu His Arg Met Asn Ala Ser Gln Met Ser			
1540	1545		1550
ggc agg gct aat ctg gat atg tgt ctg tat gac gga ggt gtg aaa gcc			6798
Gly Arg Ala Asn Leu Asp Met Cys Leu Tyr Asp Gly Gly Val Lys Ala			
1555	1560		1565
cgt tca tta cag atg aag ata gaa gga agc aat aag tca ggt acg gga			6846
Arg Ser Leu Gln Met Lys Ile Glu Gly Ser Asn Lys Ser Gly Thr Gly			
1570	1575		1580
ttt cag gtt ata aag agc gat tct gct gat acg att gat tat gcg gtc			6894
Phe Gln Val Ile Lys Ser Asp Ser Ala Asp Thr Ile Asp Tyr Ala Val			
1585	1590	1595	1600
agt atg aat tat ggg gga cga agt att cct gtc acc cgt ggc gtg gag			6942
Ser Met Asn Tyr Gly Gly Arg Ser Ile Pro Val Thr Arg Gly Val Glu			
1605	1610		1615
ttc agt ctg gat aac gtg gat aaa gca gca acg cgt ccg gtg gta ctt			6990
Phe Ser Leu Asp Asn Val Asp Lys Ala Ala Thr Arg Pro Val Val Leu			
1620	1625		1630
ccc ggg caa cgg cag gcg gta cgt tgt gtg cca gtg ccc ctt acc ctg			7038
Pro Gly Gln Arg Gln Ala Val Arg Cys Val Pro Val Pro Leu Thr Leu			
1635	1640		1645

aca aca caa ccc ttt aac atc aga gag aag cgt tct ggt gag tat cag 7086
 Thr Thr Gln Pro Phe Asn Ile Arg Glu Lys Arg Ser Gly Glu Tyr Gln
 1650 1655 1660

gga acg ctg aca gtg aca atg ctg atg gga aca caa acc ccc tga 7131
 Gly Thr Leu Thr Val Thr Met Leu Met Gly Thr Gln Thr Pro
 1665 1670 1675

cagtaattat ttatttttatt gatattctttc ttatatgggtt ttttaaatca gagttctctt 7191
 tatataacttg tttttatttaa taaagagaat ctatttcactt atgaaatca atgcgtgagg 7251

ttctgctttc ct atg act gtg tat tta gat gat aaa gat aaa gaa tta ttg 7302
 Met Thr Val Tyr Leu Asp Lys Asp Lys Glu Leu Leu
 1680 1685 1690

aaa gaa atc caa aaa gat tgt gca caa act tta tgg caa ctt gca tat 7350
 Lys Glu Ile Gln Lys Asp Cys Ala Gln Thr Leu Trp Gln Leu Ala Tyr
 1695 1700 1705

aaa gtg gga ctt acg ccc aca cca tgt ttc aaa cgt tta aaa aaa ctt 7398
 Lys Val Gly Leu Thr Pro Thr Pro Cys Phe Lys Arg Leu Lys Lys Leu
 1710 1715 1720

aaa gac agg ggg gtt atc att ggt cag ttc gct tta ttg gat aag gaa 7446
 Lys Asp Arg Gly Val Ile Ile Gly Gln Phe Ala Leu Leu Asp Lys Glu
 1725 1730 1735 1740

aaa cta ggt ctt tca ctt aat gtc ttt att atg att aac ata tct gag 7494
 Lys Leu Gly Leu Ser Leu Asn Val Phe Ile Met Ile Asn Ile Ser Glu
 1745 1750 1755

gag caa tac gct agt att tct gag aaa ata aag tca atg cct gag gtt 7542
 Glu Gln Tyr Ala Ser Ile Ser Glu Lys Ile Lys Ser Met Pro Glu Val
 1760 1765 1770

att gcc ttt tat cga att tct gga tca ttt aat tat tta atg cat aca 7590
 Ile Ala Phe Tyr Arg Ile Ser Gly Ser Phe Asn Tyr Leu Met His Thr
 1775 1780 1785

gta ttt aca gat atg aac gat tac tat agt ttt tat gag aaa ata ata 7638
 Val Phe Thr Asp Met Asn Asp Tyr Tyr Ser Phe Tyr Glu Lys Ile Ile
 1790 1795 1800

tta act aat tct tca att agt gga tct gca tcg agc ttt gtt ctt gag 7686
 Leu Thr Asn Ser Ser Ile Ser Gly Ser Ala Ser Ser Phe Val Leu Glu
 1805 1810 1815 1820

caa ata aag gaa aca aac gaa ctg tca gtg tga aagtgtgatg tgtacttact 7739
 Gln Ile Lys Glu Thr Asn Glu Leu Ser Val
 1825 1830

gatttaatac attattatcc ttcttacgga acaacaacgg cagattgcgg ctgttgaaca 7799
 aggatttttaa tcagcagtgg tgaaattaag cggcacagaa taacacagcg gaatatcaca 7859
 tggttaaata tcaccccggtg catgtaacaa aaaaccgcat taaaacagat gatgttactg 7919
 atatttattt cggtgaaccc ttctggaaaa aaggcgaaaa ccacataatt gagtcattga 7979

tgttttttga agagttacaa aagtcattta atttattcaa ccataaatat gggttaaata 8039
 aatataact caggatcccc tgggaatttg tgctacata tatggaaagg atcagtaaata 8099
 taaatagcgt cgggttattt gctgtttctg ttgacttta taacaaccac aaatttctga 8159
 gcgagtacat caggagtcgc agagattatg gtatggaagt ttggtttgat ttttgtggta 8219
 aacattctta ttccagtga attaaaaacc ttggattctt ttttcaggct tgcgtagtgc 8279
 ctctgatcc taattttatt agtagtgtt atcattatca taagttccaa aagattcttg 8339
 tcggggatat aaatgatgta gaacagaggg ccgtgtacca gaacgaagt gattacatgt 8399
 atggaatgca atggccatcg tcatatgacg gttttttctt tcgggatcat aaaaaaatg 8459
 aaacttggtg tatataacag aaggagtga aatttgaatc aaaaatatct tattttattt 8519
 ttgtttaatt attgttttgt tttttattac gattaaatat aaagaacatc attgttcgtg 8579
 cgggtggggag gctggaagt taggggatga ccgtttatca acaattttat tacagccacc 8639
 atacgaatgg ttatatatg cactagatgt attattttag tttaatatat cgatggttgc 8699
 tatttgcatt gatgatgttc cgttacatta aggaatatac atctgtatct cgttatagc 8759
 aactcacat tactaatcat tattaatatg agtgtggttc ttgttttacg catgcatggt 8819
 tgcattgtac gttaaattta aatgagctga ctgtatgaat tctaaatact ttagagaggt 8879
 gttttttgtc tcggtagtgt ttatatatt attttatttg gtgttatttg cagccagtgc 8939
 tcatgctgaa ggcggtttca gatctggagg cattgggtta tttatgacg gaacaagaga 8999
 gatgctactg tagagataat aaattctgct aaagattccc caattcttgt gcattgacat 9059
 cctccacgtc ctgaagggcg tgggttcctg ctccaacggg ctgcctgact gcacgtcct 9119
 tccacaggca agcacggcgt gtcccgctct aaaaatttac gcgcgcggtt tacatcggcg 9179
 ttgcagtat atcttcatac cagacacttg taagtatctc gcataatcgt gccattcaca 9239
 ttagagatc atac 9253

<210> 2

<211> 236

<212> PRT

<213> Salmonella typhi

<400> 2

Met	Asn	Phe	Lys	Asp	Thr	Leu	Pro	Gly	Val	Phe	Leu	Cys	Val	Ala	Met
1				5				10					15		

Phe	Ala	Cys	Gly	His	Ala	Arg	Ala	Asn	Met	Leu	Val	Tyr	Pro	Met	Ala
	20							25					30		

Ala	Glu	Ile	Asn	Ser	Ser	Arg	Glu	Glu	Ala	Thr	Ser	Leu	Phe	Val	Tyr
	35						40					45			

Ser Lys Ser Asp His Val Gln Tyr Ile Arg Thr Arg Ile Met Arg Ile
 50 55 60
 Glu His Pro Gly Met Pro Gln Glu Lys Glu Val Pro Ala Gly Asn Asp
 65 70 75 80
 Ile Glu Thr Gly Leu Val Val Ser Pro Glu Lys Phe Ala Leu Ser Pro
 85 90 95
 Gly Thr Lys Lys Thr Ile Arg Val Ile Ser Thr Gln Ala Pro Glu Arg
 100 105 110
 Glu Glu Ala Trp Arg Val Tyr Phe Glu Ala Val Pro Glu Leu Glu Asp
 115 120 125
 Asp Pro Gln Ala Gly Gly Lys Gln Asn Ser Ser Val Ser Val Asn Leu
 130 135 140
 Val Trp Gly Val Leu Leu Arg Val Ser Pro Ser Asp Pro Arg Pro Ala
 145 150 155 160
 Leu Val Thr Asp Gly His His Leu Leu Asn Thr Gly Asn Thr Arg Leu
 165 170 175
 Ser Leu Ile Arg Ala Gly Asn Cys Asp Thr Thr Cys His Trp Gln Asn
 180 185 190
 Ile Gly Lys Ser Ile Tyr Pro Gly Gly Ser Ala Asp Ile Pro Ala Gly
 195 200 205
 Ile Lys Ser Asn Ala Phe Arg Val Glu Tyr Arg Thr Gly Ala Asn Ser
 210 215 220
 Pro Val Ile Ser Ala Asp Leu Thr Ala Ala Gly Lys
 225 230 235

<210> 3
 <211> 191
 <212> PRT
 <213> Salmonella typhi

<400> 3
 Met Tyr Thr Glu Cys Thr Tyr Ile Thr Val Ile Asn Asn Lys Ala Arg
 1 5 10 15
 Leu Phe Phe Met Asn Met Lys Thr Ser Phe Ile Ala Ala Val Ala
 20 25 30
 Leu Ala Thr Val Tyr Ser Phe Ser Val Ser Ala Val Gln Lys Asp Ile
 35 40 45
 Thr Val Thr Ala Asn Ile Asp Ser Thr Leu Glu Leu Leu Gln Ala Asp
 50 55 60
 Gly Ser Ser Leu Pro Ser Thr Met Lys Leu Asp Phe Met Pro Gly Lys
 65 70 75 80
 Gly Leu Val His Lys Ser Leu Gln Thr Arg Leu Tyr Ser Asn Asp Gln
 85 90 95

Thr Lys Ser Val Asn Val Lys Leu Leu Asn Ala Pro Gln Leu Ile Asn
 100 105 110

Val Leu Asp Pro Thr Lys Thr Ile Asp Met Glu Val Thr Leu Gly Gly
 115 120 125

Arg Ser Leu Thr Thr Thr Asn Ser Val Leu Glu Ala Lys Thr Leu Phe
 130 135 140

Pro Asp Gly Lys Thr Gly Asp Ala Ser Ala Leu Leu Asn Leu Asp Ile
 145 150 155 160

Gly Gln Lys Ala Gly Ala Ala Leu Gln Asn Leu Pro Ala Gly Glu Tyr
 165 170 175

Ser Gly Leu Val Ser Leu Val Ile Ser Gln Ala Val Thr Ala Gly
 180 185 190

<210> 4

<211> 889

<212> PRT

<213> Salmonella typhi

<400> 4

Met Tyr Tyr Leu Leu Gly Leu Cys Ser Phe Thr Ser Gln Ala Thr Leu
 1 5 10 15

Ile Pro Pro Pro Gly Phe Glu Ser Leu Leu Glu Gly Gln Thr Glu Gln
 20 25 30

Ile Glu Val Leu Leu Pro Gly His Ser Leu Gly Leu Phe Pro Val Val
 35 40 45

Val Lys Pro Asp Thr Val Gln Phe Met Ser Pro Leu Met Val Leu Glu
 50 55 60

Ser Ser Gly Leu Ala Ala Leu Pro Ala Ala Glu Arg Gln Lys Ala Leu
 65 70 75 80

Ala Ala Leu Ser Arg Pro Leu Leu Arg Asn Ser Asn Leu Val Cys Gly
 85 90 95

Val Ser Glu Ala Lys Asp Ser Ser Glu Cys Gly Tyr Val Ala Thr Asp
 100 105 110

Lys Glu Asp Val Ala Val Ile Phe Asp Glu Asn Asn Ala Gln Leu Ser
 115 120 125

Leu Phe Leu Asn Arg Asp Trp Leu Pro Asp Glu Glu Arg Arg Asp Lys
 130 135 140

Arg Trp Leu Thr Pro Thr Pro Glu Gly Val Ser Ala Phe Ile His Arg
 145 150 155 160

Gln Thr Leu Tyr Leu Ser Asp Asp Leu His Ser Arg Asn Met Thr Leu
 165 170 175

Asn Gly Ser Gly Ala Leu Gly Leu Gly Asp Gly Arg Tyr Leu Gly Gly
 180 185 190
 Asp Trp Ala Ala Ile Trp Asn Gln Ser Glu His Tyr Asn Asn Ser Gln
 195 200 205
 Ala Trp Phe Asp Asn Leu Phe Val Arg Gln Asp Leu Gly Asn Gln Tyr
 210 215 220
 Tyr Leu Gln Ala Gly Arg Met Asp Gln Arg Asn Leu Ser Ser Ala Thr
 225 230 235 240
 Gly Gly Asp Phe Gly Phe Ser Leu Leu Pro Leu Ser Arg Phe Asp Gly
 245 250 255
 Leu Arg Thr Gly Thr Thr Gln Ala Tyr Val Asn His Glu Val Asp His
 260 265 270
 Asn Ala Thr Pro Val Met Val Gln Val Thr Arg Asn Ala Arg Ile Asp
 275 280 285
 Ile Tyr Arg Gly Ser Glu Leu Leu Gly Ser Gln Phe Leu Thr Pro Gly
 290 295 300
 Met His Thr Leu Asp Thr His Ser Leu Pro Pro Gly Ser Tyr Pro Leu
 305 310 315 320
 Ala Leu Arg Val Tyr Glu Asp Gly Ile Leu Arg Arg Thr Glu Thr Gln
 325 330 335
 Pro Phe Ser Lys Gly Gly Asn Ser Phe Ser Ala Gln Thr Gln Trp Phe
 340 345 350
 Ile Gln Gly Gly Leu Glu Asp Thr Gly Asp Lys Ala Ser His Tyr Asp
 355 360 365
 Gly Glu Thr Val Met Ala Ala Gly Phe Gln Thr Gly Leu Arg Lys Asn
 370 375 380
 Ile Ser Leu Thr Glu Gly Ile Ser Leu Ala His Glu Ala Trp Tyr Ser
 385 390 395 400
 Glu Thr Arg Leu Asn Ser Gln His Ala Val Leu Asp Gly Thr Leu Asp
 405 410 415
 Leu Ser Ala Gly Ile Leu His Gly Thr Asp Ser Thr Ser Gly Asn Thr
 420 425 430
 Glu Gln Val Thr Tyr Asn Asp Gly Phe Ser Ala Ser Leu Trp Arg Asn
 435 440 445
 His Thr Glu Ser Asp Ala Cys Ser Gly Arg His Pro Gln Ser Val His
 450 455 460
 Ala Ser Met Thr Cys Gln Thr Ser Met Asn Ala Ser Leu Ser Val Ser
 465 470 475 480
 Val Gly Asn Trp Tyr Ala Leu Leu Gly Tyr Ser Thr Ser Arg Thr Glu
 485 490 495

Gly Arg Pro Val Tyr Arg Gly Tyr Asp Asp Asn Ser Asp Lys Glu Asn
 500 505 510
 Val Phe Trp Arg Gln Ala Tyr Ile Pro Ala Ser His Arg Glu Ser Ala
 515 520 525
 Gln Ala Ser Ala Thr Tyr Ser Leu Asn Met Ala Gly Met Asn Ile Asn
 530 535 540
 Thr His Gly Gly Val Trp Arg Thr Arg Asn Asp Gly Val Asn Asp Asp
 545 550 555 560
 Gly Leu Phe Met Ser Val Ser Val Ser Tyr Ala Ser Gln Pro Pro Thr
 565 570 575
 Met Thr Gly Ser Asn Arg Tyr Thr Ser Ala Gly Thr Asp Ile His Ser
 580 585 590
 Ser Arg Asn Gln Lys Thr Gln Thr Ser Trp Asn Val Asn His Val Arg
 595 600 605
 Ser Trp Gln Gln Asp Leu Tyr Arg Glu Leu Ser Val Gly Phe Ser Gly
 610 615 620
 Tyr Asn Asp Asp Ser Trp Ser Gly Ser Leu Gly Gly Arg Met Ser Gly
 625 630 635 640
 Arg Met Gly Glu Leu Ser Ala Thr Ile Ser Asn Ser His Gln Arg Asn
 645 650 655
 Ala Gly Ser Ala Ser Ser Leu Thr Ala Gly Tyr Ser Ser Ser Leu Ala
 660 665 670
 Leu Ser Arg Asn Gly Leu Phe Trp Gly Gly Gly Gln Asp Gly Glu Pro
 675 680 685
 Ala Ser Gly Met Ala Val Asn Val Glu Ser Glu Gly Asp Glu Gly Ser
 690 695 700
 Ser Gly Lys Val Val Ser Val Arg Gly Ser Ser Gln Pro Phe Ser Leu
 705 710 715 720
 Gly Phe Gly Gln Gln Ser Leu Leu Leu Met Glu Gly Tyr Asn Ala Thr
 725 730 735
 Glu Val Thr Ile Glu Asp Ala Gly Val Ser Ser Gln Gly Met Ala Gly
 740 745 750
 Val Lys Ala Gly Gly Gly Ser Arg Cys Tyr Phe Leu Thr Pro Gly His
 755 760 765
 Leu Leu Val His Asn Ile Ser Ala Ser Met Ser Arg Leu Tyr Val Gly
 770 775 780
 Arg Val Leu Asp Lys Asp Gly Arg Pro Leu Leu Asp Ala Gln Pro Leu
 785 790 795 800
 Asn Tyr Pro Phe Leu Ser Leu Gly Pro Ser Gly Arg Phe Ser Leu Gln
 805 810 815

Ser Glu His Lys Glu Ser Ser Leu Trp Leu Leu Ser Lys Asn Arg Ile
820 825 830

Leu Arg Cys Pro Met Ser Val His Lys Arg Arg Asp Val Met Gln Val
835 840 845

Val Gly Asp Val Arg Cys Glu Leu Ser Asp Val Asp Ala Leu Pro Gln
850 855 860

Ala Leu Gln Ile Ser Pro Arg Val Ile Arg Leu Leu Asn Val Ala Gly
865 870 875 880

Leu Leu Arg His Ser Val Gln Glu Ala
885

<210> 5

<211> 359

<212> PRT

<213> Salmonella typhi

<400> 5

Met Ser Asn Lys Met Lys Trp Thr Ser Met Thr Ala His Trp Ser Ala
1 5 10 15

Ile Ile Asn Phe Ile Arg Lys Tyr Val Tyr Pro Ala Arg Ile Ile Ala
20 25 30

Ile Leu Leu Met Ala Gly Ala Thr Leu Pro Gln Val Ala Asp Ala Ile
35 40 45

Thr Val Asp Leu Asn Tyr Asp Lys Asn Asn Val Ala Val Ile Thr Pro
50 55 60

Val Trp Ser Gln Glu Trp Ser Val Ala Asn Val Leu Gly Gly Trp Val
65 70 75 80

Cys Arg Ser Asn Arg Asn Glu Asn Glu Gly Ala Cys Glu Glu Thr His
85 90 95

Leu Val Trp Trp Tyr Ala Phe Gly Ala Tyr Ser Lys Ile Arg Leu Arg
100 105 110

Phe Arg Glu Gln Ile Ser His Ala Glu Ile Thr Leu Ile Leu Leu Gly
115 120 125

Ser Val Arg Asp Ala Cys Tyr Thr Gly Val Ile Asn Met Asn Ala Ala
130 135 140

Ala Cys Gln Trp Gly Arg Ser Leu Lys Leu Arg Ile Pro Ser Glu Glu
145 150 155 160

Leu Ala Lys Ile Pro Thr Ser Gly Thr Trp Lys Ala Thr Leu Val Leu
165 170 175

Asp Tyr Leu Gln Trp Gly Gly Asp Asp Pro Leu Gly Thr Ser Thr Thr
180 185 190

Asp Ile Thr Leu Asn Val Thr Asp His Phe Ala Glu Asn Ala Ala Ile
195 200 205

Tyr Phe Pro Gln Phe Gly Thr Ala Thr Pro Arg Val Asp Leu Asn Leu
210 215 220

His Arg Met Asn Ala Ser Gln Met Ser Gly Arg Ala Asn Leu Asp Met
225 230 235 240

Cys Leu Tyr Asp Gly Gly Val Lys Ala Arg Ser Leu Gln Met Lys Ile
245 250 255

Glu Gly Ser Asn Lys Ser Gly Thr Gly Phe Gln Val Ile Lys Ser Asp
260 265 270

Ser Ala Asp Thr Ile Asp Tyr Ala Val Ser Met Asn Tyr Gly Gly Arg
275 280 285

Ser Ile Pro Val Thr Arg Gly Val Glu Phe Ser Leu Asp Asn Val Asp
290 295 300

Lys Ala Ala Thr Arg Pro Val Val Leu Pro Gly Gln Arg Gln Ala Val
305 310 315 320

Arg Cys Val Pro Val Pro Leu Thr Leu Thr Thr Gln Pro Phe Asn Ile
325 330 335

Arg Glu Lys Arg Ser Gly Glu Tyr Gln Gly Thr Leu Thr Val Thr Met
340 345 350

Leu Met Gly Thr Gln Thr Pro
355

<210> 6

<211> 151

<212> PRT

<213> Salmonella typhi

<400> 6

Met Thr Val Tyr Leu Asp Asp Lys Asp Lys Glu Leu Leu Lys Glu Ile
1 5 10 15

Gln Lys Asp Cys Ala Gln Thr Leu Trp Gln Leu Ala Tyr Lys Val Gly
20 25 30

Leu Thr Pro Thr Pro Cys Phe Lys Arg Leu Lys Lys Leu Lys Asp Arg
35 40 45

Gly Val Ile Ile Gly Gln Phe Ala Leu Leu Asp Lys Glu Lys Leu Gly
50 55 60

Leu Ser Leu Asn Val Phe Ile Met Ile Asn Ile Ser Glu Glu Gln Tyr
65 70 75 80

Ala Ser Ile Ser Glu Lys Ile Lys Ser Met Pro Glu Val Ile Ala Phe
85 90 95

Tyr Arg Ile Ser Gly Ser Phe Asn Tyr Leu Met His Thr Val Phe Thr
100 105 110

Asp Met Asn Asp Tyr Tyr Ser Phe Tyr Glu Lys Ile Ile Leu Thr Asn
115 120 125

Ser Ser Ile Ser Gly Ser Ala Ser Ser Phe Val Leu Glu Gln Ile Lys
130 135 140

Glu Thr Asn Glu Leu Ser Val
145 150

References:

- American Institute of Medicine. (1986) New vaccine development: establishing priorities. Washington, DC: National Academy Press.
- 5 Ashcroft, M. T., Ritchie, J. M., Nicholson, C. C. (1964) *Am. J. Hyg.* 79:196-206.
- Levine, M. M., Taylor, D. N., Ferreccio, C. (1989) *Pediat. Infect. Dis. J.*, 8:374.
- 10 Popoff, M. Y. & Le Minor, L. (1992) *Antigenic formulas of the Salmonella serovars* (WHO Collaborating Center for Reference and Research on *Salmonella*, Institute Pasteur, Paris).
- Reeves, M. W., Evins, G. M., Heiba, A. A., Plikaytis, B. D. & Farmer III, J. J.
15 (1989) *J. Clin. Microbiol.* **27**, 313-320.
- Rowe, B., Ward, L.R., and Threlfall, E.J. (1997) *Clinical Infectious Disease* 24:(Suppl 1) S106-9
- 20 Salyers, A. A. & Whitt, D. D. (1994) *Bacterial Pathogenesis: A molecular approach.* (ASM Press, Washington D.C.).
- Yugoslav Typhoid Comission. (1964) *Bull. WHO* 30:623-30.